Going the distance

Ruhland, Hendrik Georg

Going the distance: A Non-Linear Approach To Change In Language Development

H.G. Ruhland

Cover design by Graphic designers studio Lynck & Zoon in cooperation with Rick Ruhland

Printed by Print Partners Ipskamp, Enschede.

The research presented in this thesis was carried out at the Department of Developmental and Experimental Clinical Psychology of the Heymans Institute, University of Groningen, The Netherlands, with a grant (grant number: 575 - 07 - 020) of the Foundation for Behavioural Sciences (SGW) funded by the Netherlands Organisation for Scientific Research (NWO).

RIJKSUNIVERSITEIT GRONINGEN

Going the distance

A Non-Linear Approach To Change In Language Development

PROEFSCHRIFT

ter verkrijging van het doctoraat in de
Psychologische, Pedagogische en Sociologische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Van der Woude
in het openbaar te verdedigen op

donderdag 19 maart 1998
des namiddags te 1.15 uur

door

Hendrik Georg Ruhland

geboren op 18 april 1966

te Zwolle

Promotor: Prof. Dr. P.L.C. van Geert

Acknowledgement

Catastrophes happen, but luckily, the good things in life help to make it through the daily bustle. Music, for example, is very important to me. That becomes clear if one looks at the chapter titles in this study: they are derived from songs or album titles of favourite singers and bands. Without music, humour, the arts, food (I really like to cook and eat, and it shows), sports, and much more, life would not be half as pleasant. To give a very good example: one of the most humorous conversations comes from the movie *Airplane!* (aka *Flying High*) (1980).

> Doctor Rumack: Captain, how soon can you land?
> Ted Striker: I can't tell.
> Doctor Rumack: You can tell me, I'm a doctor.
> Ted Striker: No, I mean, I am just not sure.
> Doctor Rumack: Well, can't you take a guess?
> Ted Striker: Well, not for another two hours.
> Doctor Rumack: You can't take a guess for another two hours?

Nice conclusion by a doctor.

But the greatest of joy is brought to me by places that I call mystical outposts. These places, Scotland and Ireland in particular, are definitely the places I go to on holiday. Nobody seems to understand why, not even I, but those are the places where past, present and future slide into one, leaving behind an incredible peace of mind and idleness. In the countries of stone circles, cairns, folk music, whisk(e)y, etc. etc., the pace of the rush hour is set by the speed of sheep chewing grass.

Anyway, change and variation makes life worthwhile. This becomes clear in my ‘scientific history’: my Master’s thesis was on the acquisition of negative polarity items, my PhD thesis is on sudden change in development. This change of subject is quite ‘me’. Never a dull moment.

Here, I would like to thank, first of all, Prof. Paul van Geert. We met nearly six years ago when I wanted to cross my old boundaries, and to seek new ones. We have rather different backgrounds. He is a developmental psychologist interested in language, I am a linguist interested in psychology. He told me about psychological problems and issues of (language) development, and I told him a bit about linguistics. Paul, thank you for these five years of sharing knowledge, playing with language, and exchanging ‘a burst of laughter’ every fifteen minutes (on average).

The Committee that read the thesis consisted of Prof. Frans Zwarts, Prof. Brian Hopkins, and Prof. Lex Kalverboer, and I thank them for taking part in the committee. Dear Frans, I graduated in 1991 on a subject which is still one of your research topics, namely negative polarity items (NPI). I very much enjoyed doing my masters on the acquisition of NPI’s. I remember your broad views on language and language development. In my second year, you argued, in a lecture on Semantics that the study of language needs psychology, philosophy and neurology (amongst others) to fully comprehend the nature of language. You always have a big smile on your face. Dear Brian, when I was in Lancaster last year (in August), I was not aware of you becoming a member of the committee. I only knew you as a member of the Transitions group. Somehow, I thought you were not interested in research, the study of development, or the social sciences in general, but in the knowledge that results from research. Enjoying life whilst smoking a cigar and having fun with other people is the picture I have in mind. Dear Lex, I still see you as a classic professor in the sense that you are a homo universalis. Apart from research, you are a very good ice skater (;-)), you make wonderful paintings, you sing classical music, and if you had wanted another job, you might be a cabaret artist.

In the past five years, I have been a member of a group of PhD students and senior researchers. Together, we were the 'Aandachtsgebied Ontwikkelingstransities', the Developmental Transitions Group. I like to thank Carolina de Weerth, Pascal Hartelman, Phillipa Butcher, Raymond Wimmers, Brian Hopkins, Han van der Maas, Lex Kalverboer, Peter Molenaar, Geert Savelsbergh and, of course, Paul van Geert, and the visiting professors Ron Barr and Kurt Fischer, for the inspiring hours we spent on discussing the application of non-linear models and theories like Catastrophe theory to ontogenetic human development.

I also like to thank the group of researchers and the secretary of the Department of Developmental and Experimental Clinical Psychology for shedding a new light on science in general and the development of us humans in particular. I especially want to mention Saskia Kunnen, Martin Cats, Edith Heurkens and (just before the 'last orders, ladies and gentlemen') Marijn van Dijk and Martine Verheul, with whom I shared the room for four years or less. Furthermore, I express my thanks to Herbert Hoytink who helped me out with statistics in the last phase of the research. Likewise, I would like to thank the "old home of linguistics": the Department of Dutch Literature and Linguistics and the General Linguistics Department. I especially would like to thank Jan Wouter Zwart who helped me out with my intuitions on linguistic issues. It felt good to come back every now and then to the scientific nest I was raised. My thanks go the students, especially Dieuwke de Goede, for reflecting on my 'linguistic' thoughts on development. I furthermore would like to thank Gerard Bol, Caroline Elskamp, Puck Goosens, Evelien Krikhaar, Paulien Rijkhoek, Frank Wijnen, and several students who did a labourious job: they recorded the children and transcribed over a hundred hours of child language. Furthermore, I would like to thank Frank Wijnen who, as a psychologist, has helped me understanding the sometimes subtle differences between a psychological and a linguistic approach to language development. Frank, thanks for your creative and broad-minded views, that fed our discussions on language, psychology and (sometimes) life. Lastly, I owe much to the BCN, the graduate school for the study of Behavioral, Cognitive and Neural Sciences. My special thanks go to Mark Kas, the coordinator of the BCN. Maybe we can go to Milan again someday.

I like to write, and I do it everyday, but it is difficult to express one's ideas in a very precise way in a language that is not the native language. I would like to thank Saskia Kunnen for her comments on the reasoning and structuring of the text, and especially Bas van Mierlo for his 'miereneukerig' (miereNneukerig), i.e. close, reading of the thesis. And if you close this book, you will see the cover that was made in cooperation with Maaike Hamelynck from the graphic designers studio Lynck & Zoon. Maaik, thanks for being a friend and helping me out with the cover.

Life cannot be without friends, and with them I share my thoughts, hopes, fears and pleasures. Adriaan (Let's *2 voor 12* again), Annet, Bas, Coen, Dina, Frans, Janette, Maaike, Mirjam, Oene, Paul, and Roel, thank you for being around. However, I must name two in particular: Bas van Mierlo and Roel Jonkers, for being my accompanying ‘penguins’. In addition Bram ‘told’ me most about language development: I recorded his language every two weeks for nearly two years, and even though I never used his data, he succeeded in what theories and models cannot do: show what a child does in real life.

Love is surely all that matters. Martine, thank you for giving me, everyday, a cup of love potion no. 9.

Without making a pun, my parents and I experienced a catastrophe, to say the least. Their youngest son, my “little brother”, died unexpectedly in June 1996, at the age of 27. Their sadness as well as their support and their will to move on made me decide to finish this thesis, although I have struggled with the thought of throwing it all away. Mum, dad, this is also for you.

To the memory of my brother Luc (June 13th, 1969 - June 13th, 1996)

Contents

Page

1

Begin the begin: the introduction

Let's begin again, begin the begin, Let's begin again like Martin Luther Zen
The mythology begins the begin Answer me a question I can't itemize

Begin the Begin by **REM** from the CD **Life Rich Pageant**

1.0 The Droste effect

Causes of change have traditionally been written in formulas or equations like "If y happens, then x is the cause." These sorts of rather simple, mostly *linear* relationships are still in use to describe or explain all kinds of behaviour. However, in recent years, new research is focussing on more complex relationships. For example, dynamic systems theory (see also chapter 3) does not use linear, but non-linear relationships. These relationships lead to fascinating new theoretical and practical insights. To give an example: in the field of these non-linear theories researchers discovered the *fractal*. A fractal is a mathematical structure. Its overall structure is infinitely repeated on smaller scales. Usually, fractals are depicted as a drawing, and they are supposed to be very common. For example, the drawings found on tins of the Dutch chocolate manufacturer Droste, hence the Droste effect, shows the effect of a drawing in a drawing on smaller and smaller scales. Fractals have become popular with the introduction of the so called Chaos theory. This theory was soon to be followed by other non-linear theories like, for instance, Catastrophe theory. These non-linear theories constitute an important part of this thesis, because linear models fail to explain at least some of the models in developmental psychology.

The central research question of this dissertation, formulated in the next section, pays attention to both linguistic and psychological theories of human development in combination with empirical evidence. Language development is analysed by applying non-linear theories to quantified language variables to find evidence for discontinuities in human development. Begin the begin, so to speak a very small fractal captured by the rock song by REM, refers to the start of human development, to the way a child develops, and to how to define the start of acquiring his[1] language.

1.1 A brief history of views on change

The study of change in the development of any (including human) behaviour is not new. In the course of history, there has always been a discussion whether nature is static or ever changing. In the 6th and 5th century BC, the Greek were concerned with change in nature, and they had two opposite views. Parmenides, on the one hand, had a rather conservative view on the world: everything in our world had existed ever since the world was created, no change has ever occurred. According to Heraclite, on the other hand, everything flows, and change is essential to nature (Störig, 1972). In his *Historia Animalium*, Aristotle takes yet another point of view. In book VIII, after having discussed the physical characteristics of animals, he states that *Natura non facit saltus* ('nature does not jump'). A problem was added to the discussion on the nature of men by Plato. In *Meno*, Plato worked out some of his ideas on the problem of novelty. Is everything we humans are innate or is everything we are handed over by the environment (e.g. our parents)? Plato's choice for innateness and his ideas on change have been challenged. This dissertation illustrates that the ancient discussion on human behaviour and change has not ended at the end of this millennium.

The question of change or no change still exists in our time. Views on the development of language, with respect to change and statis, were challenged in the late fifties and early sixties when a behaviourist psychologist, Skinner, and a young linguist, Chomsky, set fire to the views on language and to the issues of change and novelty in language development. Chomsky stressed the importance of *structure* in language. It was recognised that the structure and order in language development were by no means accidental.

The criticism by Chomsky and other generative grammarians was directed against research preceding the introduction of generative grammar. Language, in those pre-generative days, was seen as a collection of words, a way to communicate, a carriage to express thoughts. There was a strong influence from philosophy and from (developmental) psychology. Emphasis was put on, for example, philosophical problems. However, since there is *structure* in language (expressed in the grammar of a language), the question is how this structure is acquired in the process of language development. One of the goals of this study is to explain the development of these structures (e.g.

[1] *He* is used as a gender neutral pronoun to refer to a child in general, male or female.

grammar), development is seen in its context of a process, i.e. something that has an index of time, and it's main characteristic is *change*.

The forms of change, the relationship between these forms, and the (underlying) reasons and causes of change are the main topics of this dissertation. The ancient ideas about the nature of change still stand, but new methods and tools have been developed to demonstrate all sorts of change, including gradual, fast, and sudden discontinuous change. These tools and methods will be applied to developmental issues to answer the research questions in the next section.

1.2 'Houston, we have had a problem'

When Apollo 13, one of the space rockets that went to the moon and back, got into trouble, all sails had to be set to get the astronauts back to Earth safely. Due to an unforeseen problem the Apollo flight could not return to earth smoothly using standard procedures. Inventive minds had to come up with a creative solution. A non-standard situation like the problem with Apollo 13 calls for a non-standard solution. We will see that a non-standard approach is also needed for at least some of the questions in developmental psychology.

In the study of development, some approaches assume a linear or continuous underlying change of psychological processes, which is reflected in the use of statistical tests like regression. Other approaches, however, claim that development is discontinuous (e.g. in Piagetian theory where stages are assumed). The problem is how to prove discontinuities and how to define stages. Various scholars have tried to provide a solution to the problem (see for an overview Boom, 1993). The many (near-)synonyms of stage (like period, phase and level) are an indication of more problems, since the term 'stage' embraces two concepts, namely equilibrium and instability. These concepts reveal new problems: how are equilibrium and instability defined? No satisfactory answers have been given so far. So, in developmental psychology, 'we have had a problem'. There is a claim of discontinuity, but there is no model for explaining discontinuities. This is a surprise, since the assumption of discontinuous development is widespread within, for example, (neo-)Piagetian research. So, we need to come up with a non-standard solution, since standard solutions refer to continuous change.

Mathematical theories and methods elicit the sort of change. On the one hand, there exist, apart from linear models, growth models. These models are used to explain continuous change in development. On the other hand, there is Catastrophe theory, that is very helpful for fitting and explaining discontinuous change in development. These non-linear models have been chosen to define (dis)continuity and to answer the following questions (with the chapters in brackets):

1. What is development, and does behavioural development (i.e. language development) consist of sudden discontinuous change, or is it continuous? [Chapter 2].
2. What theories are useful to prove (dis)continuous change in development? [Chapter 3].
3. What method is needed for the longitudinal study of discontinuities, and what is the empirical evidence for discontinuous development? [Chapters 4 and 5].
4. What does discontinuity mean for the study of development in general, and for the study of language development in particular? [Chapter 6].
5. What is the relationship of the data with linguistic and psychological theories, and what do non-linear models have to offer? [Chapter 6].

The central question is:

> What are the paths of language development (in quantitative terms), and how can these paths of development be explained by linguistic and/or psychological theories?

More specific, this study reveals possible evidence for a transitional (for the time being defined as discontinuous) change in the ontogenetic (i.e. in a child) development of language.

1.3 The dissertation from a bird's-eye view

In Chapter 2, I deal with the two central concepts of this study: development and language. Development is closely related to novelty in behaviour and to the nature-nurture debate. Furthermore, I discuss the notion of discontinuities in development. In the second half of the chapter I stress that linguistic theories use (underlying) structures in language to explain language properties. Some problems and questions remain, because until recently there were no theories or models to empirically prove discontinuous behaviour.

Chapter 3 contains a discussion on a branch of mathematics that has been developed in the past twenty years. Non-linear models and theories, especially dynamic models and catastrophe theory, offer a method to study non-linear and even discontinuous change. This chapter also presents some promising studies in the field of developmental psychology.

Of course, theoretic reasoning needs empirical support, which in turn cannot do without a method. In Chapter 4 the variable of the analyses, *function words*, is discussed. The choice is based on findings by Roger Brown (1973). I also discuss the conditions that must be met to study quantitative change in language development, and I will present the subjects in this study (i.e. six children), the recordings and the transcripts employed.

In Chapter 5, the Results, I first discuss indicators from Catastrophe theory (i.e. the flags) that were found in the language of the six children. Second, the results from continuous models are presented, namely fits of the time series. This chapter ends with explanations for the change found in the data.

Chapter 6, the Discussion, starts with a summary of the evidence presented in chapter 5, and I discuss what this evidence means for the study of change in (language) development. Psychological and linguistic theories of chapter 2 return in this chapter. The question is how these theories relate to the findings from this study. I discuss non-linear models and theories like Catastrophe theory and Connectionism and I also go into the matter of the problems that come with the method of this study. Some solutions are proposed, and the completion of this thesis includes some tentative predictions based on this study, and future improvements to find more evidence for either continuous or transitional change.

2

Everything must go: language development

And if you need an explanation, Everything must go

Everything must go by **The Manic Street Preachers** from the CD **Everything must go**

2.0 Introduction

The study of *language development*[1] has created a massive body of literature, most of which is strongly influenced by linguistic theories. The words language and development, however, have several interpretations, that come with a discussion on *origin* and *novelty* in development. In the nature-nurture debate, which is dedicated to the origin of behaviour, the question is: is the behaviour of humans determined by nature or by nurture? The extreme positions are that human behaviour is either the result of genes or neurons (i.e. nothing is added from the outside and man is an *output device*), or the sole product of nurture (i.e. humans are nothing more than *input machines*). The question thus is how new behaviour can arise. This question has its roots in biology, especially in developmental biology. If all humans are but one cell at the start of conception, how do we become differentiated and how do we end up being a human with many cells, functions and all kinds of behaviour? In short, the question is whether or not humans are confined to a predetermined path of development.

Another debate, i.e. the debate on discontinuity and continuity, deals with the question whether development proceeds little by little in an Aristotelian, continuous way, without any structural reorganisations or qualitative changes, or whether development is not continuous. This issue is related to change and it is also the central theme of this thesis. In the first section of this chapter, literature is reviewed with respect to *development* and *(dis)continuity*.

The second part of this chapter (section 2.2) is dedicated to *language*. The study of language and its structure received a push forward in a new direction in the second half of this century. Theories on the structure (i.e. syntax) of language and of language

[1] The term language development is used for the discussion on language (structures) in general. The more specific term grammatical development is used for the development of sentence structures.

development are reviewed, with special attention to transitional phenomena that are predicted by these theories.

Section 2.3 contains a summary of the chapter. It states the core problems of the study discontinuous behaviour, and a first conceptualisation of the central research questions of this study.

2.1 Discontinuity in development

Introduction

The debate on the origin of development has been called the debate on **nativism** (from the Latin *natal*, meaning inborn) and **empirism** (from the Latin (for Greek) *en* (in) *peirao* (try)), and it constitutes an important part of developmental research. Apart from origin issues, the issue of novelty in behaviour (e.g. language) is theoretically very important. It addresses the question: where does (new) behaviour come from? Origin and novelty will be discussed shortly and they introduce the main subsection on discontinuity.

Development is not an unambiguous concept: there are several meanings (or definitions) of *development*[2]. Development means *growth*, and growth is defined as an increase in size or value. This means that something has been present all the time, nothing new is added. Development also means *evolution*, which derives from the Latin *evolve* meaning to unfold or to open out. *Stage of advancement* is listed as a definition of development, advancement being derived from the word *advance*, which means (among others) to rise or to move forward. If change (e.g. to increase, to open out, to move forward) is the core aspect of development, the next step is to determine the treatment of change (development) in theories on development.

Theories on development

“The primary goal of a developmental analysis is the study of change, not statis.” Riegel (1975)

Riegel’s statement on the meaning of the concept of *development* is by no means an incorrect view, it is, however, inspired by just one of the many meanings of development. The important issue is that all meanings of development assume change. *Change* is regarded as the main characteristic of development.

[2] The dictionary used here was the Concise Oxford Dictionary.

The dictionary descriptions, like other levels of analyses of development and its meaning (cf. van Geert, 1986b, 1986c), do not tell us much about the *process*, which development is after all. According to Hoppe-Graf (1989, 2), there is not even consensus on definitions of development. Despite this perhaps awkward situation, a number of developmental theories have been put forward, and they all have to account for change in behaviour, change just meaning “going from nothing to something” (e.g. from being a baby without a language or grammar to an adult with full communicative capacities and linguistic competence). These theories also have to account for the development of a subclass of behaviour, namely language.

Theories on development can be divided into two general types[3]. First, *non-stage* theories are theories that predict gradual or linear change. Skinner’s theory, for example, is a learning theory, and most learning theories by definition predict gradual or linear, or continuous change (Wanner & Gleitman, 1982). Unless the environment changes suddenly, there is a chance of sudden change in development. Take, for example, the big prehistoric dinosaurs. They became extinct when the earth was struck by a meteor and the conditions on earth were no longer viable. Second, *stage* theories predict sudden changes in development (e.g. Piaget’s equilibrium model). They incorporate by definition qualitative reorganisation, which is also known as *discontinuity*. The discussion on the exact meaning or definition of (dis)continuity, captured in a large body of research, is used for the reinterpretation of the research question. It is not my intention to force all developmental theories in this stage vs. non-stage dichotomy, but it is a useful way of categorising theories within the context of this thesis. It also helps to limit the range of explanations of the form and the mechanisms of change. Why and how does a child go from “nothing” (i.e. the newborn state) to “everything” (i.e. the adult state)? And if there are stages, why are they ordered the way they are? The answers to the questions will show us the paths of development, and they are the main subject of this book.

The origin of behaviour: innate or not?

Since general developmental theories are not theories about language development, one of the main objections against them is that they leave aside the fact, for example, that all languages have a structure (usually referred to as the *grammar* of a language), and that languages share common features (e.g. most languages use verbs to denote actions). This emergence of structures (e.g. a grammar) in language is not explained by these theories. Furthermore, most theories on psychological development assume that learning is a very

[3] See, for example, Mussen (1983) or Miller (1993) for an overview of developmental theories.

important aspect of language development. However, this learning cannot explain how children acquire the application of rules and exceptions of, for example, the use of determiners (e.g. articles) or prepositions, the mistakes a child makes, and the deduction of rules in language development.

The origin of behaviour, i.e. the important question of how any behaviour can arise, and the structure of change is determined by the starting values of the brain. These starting values have two ‘extreme’ options. Humans are completely “empty-headed” at birth, which has been called a *tabula rasa*, or everything is available to a child at birth. The second option means that humans are born with full capacities to acquire any behaviour, including language. This so-called *innateness* could explain the speed of development. If it is assumed that all language faculties are innate, then language possibly develops much faster since the ability of using language does not need time to mature (see below for maturational accounts of language development). Plato was probably one of the first to notice that the origin of thought (including language) either needs a blueprint, or is completely caused by imitation. In *Meno*, he tries to prove that man cannot be without inborn capacities. Knowledge and behaviour are too complex and too broad to be learned in a lifetime. In the same way that Plato argued, modern *nativist* theories assume a blueprint that governs development. Chomsky wondered how humans come to rich and specific knowledge, or to intricate systems of belief and understanding, when the available evidence is so poor (although it is questioned whether the evidence is indeed so meagre; cf. Bowerman (1983) for a discussion on negative evidence, i.e. the fact that children have little evidence of incorrect, ungrammatical sentences). If all behaviour is determined by birth, change depends on the blueprint (e.g. everything develops at once on an underlying level, although factors like memory might hamper the immediate execution of the blueprint).

The specific innateness approach to (language) development called *modularity* is best known from the work of Fodor (1983). The modularity discussion has been one of the inspiring discussions on how humans acquire a language. Proponrent of modularity believe in independent autonomous subsystems in psychological processes. These subsystems called modules explain the process of acquisition. According to Fodor (1983) a modular theory consists of input systems and of central systems. Input systems are autonomous from birth. There is some criticism (cf. Karmiloff-Smith, 1992) against such an approach. It holds that development is a process that goes beyond modularity. It is assumed that this process is not strongly restricted by clearly defined modules, since this would imply the absence of a creative process, which (language) development seems to be, according to developmental facts (see for an overview O’Grady, 1997). Neither strong nor weak modularity is favoured because it is unlikely to be correct: it lacks a developmental component (Karmiloff-Smith, 1992).

In this study modules are used in a linguistic way. Such modules (e.g. syntax) help understand and limit the number of possible structures that exist in a language (see also section 2.2). Furthermore, it is highly unlikely that there are no interactions between modules from birth on. This has been shown by, for example, Donahue (1986) who found phonological constraints on the development of utterances. Nevertheless, it is useful to assume modularity on a descriptive level. That is, the assumption of modules like syntax and phonology help understand the relatively independent acquisition of, for instance, phonological features or syntactic rules. On a developmental (process) level, the interaction between several modules is probably the essential motor of creative use of language. Therefore, I assume that development is modular, but that there is a strong interaction between subsystems. In development, these subsystems can be separated to study one aspect, but there is a strong interaction between the modules (see also chapter 6 for an interpretation of the results). Atkinson (1986) summarizes the modularity and innateness discussion, and he argues that this discussion has two possible outcomes. Language acquisition is either a matter of learning, in which case there is hypothesis formulating and testing, or it is a matter of language structures being *triggered* by the environment. If it is assumed that there are modules (highly organised, fairly independent subsystems), like the syntax or semantics of a language, it is conceivable that these modules rest on innate structures that may be highly specialised in terms of a blueprint, and that are triggered by the environment. It is even plausible that the language system itself, in conjunction with the input from the environment, triggers new modules in development. In chapter 6, I argue for a ‘triggering’ approach to language development .

In *learning* theories it is proposed that every kind of behaviour is caused by learning processes. In adapting to the environment, a child copies or imitates behaviour and internalizes it. In Skinner’s behaviourist theory, for instance, new behaviour is added to the existing set of behaviour by of stimulus-response chains. Criticism by Chomsky and others holds that certain regularities in mistakes in behaviour cannot be learned, since these are not present in the environment. Furthermore, the “poverty of the stimulus”-argument holds that a child receives little information about ill-formed sentences. The child does not know a priori which strings are correct in a language. Since parents make false starts, or even ungrammatical sentences, children should make the same mistakes in their language (which they do not make). Another weak point is that although no innate principles are assumed in learning theories, a child must have a capacity to learn, and since these learning principles are not explicitly taught, they must be innate. So, instead of going around the problem of innateness, the innateness assumption also plays an important part in learning theories.

Maturational theories assume that behaviour unwraps. There is a blueprint at birth, in which physical as well as psychological development is embedded. The child's abilities are subject to the laws of slow and gradual growth. The direction of development is controlled by an intrinsic mechanism. Changes in an organism or system depend on the blueprint, and new behaviour is bootstrapped by the preceding behaviour. An important concept is the critical period: biological structures determine ontogenetically developing cognitive structures. This critical period also holds that there is a limit on the age that certain types of knowledge, e.g. language, are learnable. However, maturation and critical periods cannot explain the universals that exist in the ontogenetic development of language (De Graaf & Breeuwsma, 1996). Furthermore, the 'function follows structure'-argument of maturational theories does not hold. Thelen and Ulrich (1991), for example, have shown that walking patterns are present from birth, but due to physiological constraints (e.g. muscle strength and fat proportions) these patterns remain unused until the first birthday. The 'maturational' predictions of change are that there is slow and gradual growth, since this is the sort of change we might aspect according to maturational theories (see Gibson & Petersen, 1991 for an overview of maturation and cognition).

In sum, the theories above support either inborn capacities (i.e. innateness), or capacities, learned from environment. This dichotomy is called the nativism-empirism debate, that in its pure form no longer exists in the literature on development. Usually, the best bits of both worlds are chosen to explain the nature of development. These interactionist theories of development are a combination of considerations of both nature **and** nurture, of both inborn capacities and environmental input. It is assumed that development is the result of inherent elements of a (developing) system[4].

Discontinuity and continuity in development

Introduction

It is often questioned whether or not development, more specifically cognitive development, is a continuous process, or that discontinuity is its central characteristic. Closely related to (dis)continuity are predictions about the form of changes over time. The form of change can be specified by the scores of a variable on, for instance, a test,

[4] A study of novelty can be found in De Graaf (1997).

repeated over time. In many traditional statistical tests (e.g. linear regression) it is assumed that this change is *linear* over time. Another assumption is that changes are *unimodally* distributed over groups. The assumption is that there is one mean, and all scores are distributed around this mean. Directly related to discontinuity are also all sorts of *stage*-concepts. Stage-wise development (e.g. in Piagetian theory) is in fact another conceptualization of discontinuous development.

Discontinuous change: definitions

The meanings of (dis)continuity differ across the disciplines that are relevant to this thesis (i.e. they are different in psychology, linguistics, and mathematics). Furthermore, (dis)continuity has several meanings within the field of psychology. According to Ford and Lerner (1992, 19), discontinuity refers to different ways and causes of development. Sternberg and Okagaki (1989) give two definitions of (dis)continuity. First, discontinuity is the lack of smoothness in growth curves. Second, discontinuity is the lack of constancy of the rank orders of individual differences across ages. According to Ende and Harmon (1984b), continuities refer to connectedness in development, to a linkage of early behaviour to later behaviour. Discontinuous changes are what Connell and Furman (1984) call structural or qualitative changes. In sum, there is a wide range of meanings of discontinuities. These meanings are listed below, and they summarize other meanings:

1. A phenomenon has different causes;
2. A lack of smoothness in growth curves;
3. A lack of constancy in rank orders of individual differences across ages;
4. No linkage of early behaviour with later behaviour;
5. Qualitative or structural changes.

These meanings of discontinuity lack a clear and functional application to development. For example, what does lack of smoothness mean? Or how is the absence of linkage with earlier behaviour measured? I postpone this problem for the moment, and return to the matter at the end of this section.

The second issue to be discussed is the relationship between learning theories and stage theories on the one hand and (dis)continuity on the other hand. If development is based on learning, then this process is at bottom continuous (Gleitman & Wanner, 1988). However, the opposite does not necessary hold since the outcome of an innate assumption interacts with (output) constraints. That is, if development is not learning, then it does not need to be discontinuous. If a child acquires a rule of some cognitive structure, then the execution of the rule could be hampered or limited by, for instance, memory constraints, or by other cognitive domains (e.g. learning to read the time depends on knowledge of numbers). These constraints slow down the execution of a rule. The crucial point here is that a theory that is adopted to explain a phenomenon has serious implications for the prediction of the change (of that phenomenon).

The third question is whether development is an **either-or** situation, or an **and-and** situation. Sternberg and Okagaki (1989), for example, argue that development is a process of both continuity and discontinuity. In most developmental theories it is assumed that development is either continuous or discontinuous. This is not likely to be so since development has its discontinuous processes (e.g. conservation; Van der Maas, 1993) and continuous processes that are an addition of (combining types of) knowledge (e.g. the knowlegde to count). The only way to make plausible that development is both discontinuous and continuous is the proof of empirical data. Thatcher (1991) holds that it is an unresolved issue whether human cognitive development occurs as a continuous or a discontinuous function of age. Both processes coexist and operate throughout the life-span and those processes are a common feature of postnatal human cerebral development.

However, the discussion has not been settled yet. Discontinuity refers to qualitative change, but, as Fischer pointed out, this is not sufficient (Fischer, 1983a). One of the problems is the determination of discontinuities. Fischer, Pipp and Bullock (1984) developed two tools (the scalogram and the method of multiple tasks) for measuring discontinuities in cognitive development. According to Fischer et al. (1984) the (non-)existence of discontinuities and transitions depend on scale and method of measurement. The dependence on scale is expressed in a scalogram analysis. This analysis provides the scale that is required to find developmental spurts (i.e. a feature of discontinuities). With respect to the method, they put forward that “all that is required to test discontinuity is that the distance between steps not to be confounded with the points of discontinuity” (p. 102/103). This method produces evidence for a discontinuity in terms of going from zero performance to full or nearly performance. In addition to the method of measurement of Fischer et al., Bower (1983) supposes that discontinuity can be demonstrated if there is a gap between test scores. In sum, the measurement of discontinuity in development is strongly based on the speed and the distance between changes in growth curves. This, however, is not enough to prove discontinuities.

Discontinuous change: growth curves

The discussion of (dis)continuity can be related to the first definition of Sternberg and Okagaki (1989) using growth curves. Growth curves are scores of a variable (e.g. a score on a test or the frequency of a variable in a natural situation), measured over time. The sort of change in such growth curves is either continuous or discontinuous, or asymptotic or exponential. The examples in Figure 2.1 are continuous asymptotic or continuous exponential, the examples in Figure 2.2 are discontinuous (see chapter 3 for more details on the cusp).

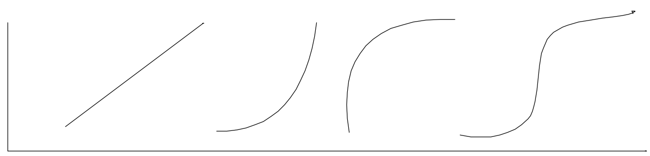


Figure 2.1. From left to right: linear, exponential, asymptotic and s-shaped growth.

Figure 2.2. Two examples of discontinuous change: discrete step (left) and the cusp (right).

In psychological research a linear increase over time is more often than not the underlying continuous model of development, but there are other elementary types of change over time. Exponential change is characterised by a slow start and a gradual infinite increase. Logistic (also called gradual or s-shaped) change consists of a slow start, after which the graph shows an increasing steepness (the tangent increases in value), after which the growth levels off to a horizontal line. Asymptotic change is when change starts of very rapidly, but decreases when it reaches a maximum (a ceiling effect).

The two graphs in figure 2.2 depict sudden change. This change starts off like gradual change (zero score, thus a horizontal line), but instead of a smooth increase the graph shows a sudden change. This lack of smoothness of the curve relates to the Sternberg and Ogasaki definition of discontinuous change.

A gradual change may look like a sudden change. That is, when the steepness of the s-shaped graph (between the horizontal start and end) increases, figure 2.1 starts to look like the discrete step in figure 2.2. It is difficult, if possible at all, to decide where the border between both sorts of change is (i.e. in a data-driven approach).

In sum, I distinguished six forms of change in time series. There is change which has no end point, i.e. linear and exponential change, and there is change with an end point, i.e. logistic, asymptotic, discrete and cusp (See chapter 3 for an extensive discussion on change in the form of a cusp). This end point is constant with respect to the score on the y-axis. The difference between logistic/asymptotic and discrete/cusp (all four have an end state) is the suddenness of change. The difference between logistic and asymptotic growth is the starting point. The difference between the discrete step and the cusp is that the cusp has an overlap of states and more than one possible jump from one state to the other. Despite the intuitively appealing descriptions, all forms of growth lack a formal criterion (e.g. how can one distinguish between a rapid gradual change and a discrete step?). Apart from a conclusive definition, the descriptions are also problematic in the sense of testing. Therefore, I present a mathematical approach to quantitative development. Non-linear models and theories are introduced in the next chapter. These models and theories allow for an estimation of parameters and state the relationship between x (a time index) and y (the variable) values in terms of a growth indication (e.g. in terms of continuity and discontinuity).

Preliminary conclusions on the concept of discontinuity

Three elements have been discussed with respect to the subject of discontinuity. First, there are many definitions of (dis)continuity. Of these definitions, I derive the following definition:

> There is a qualitative (structural) change in development in such a manner that there is no structural linkage (i.e. after the change there is a structure that was not there before) with earlier behaviour, and this change is reflected by the lack of smoothness of a growth curve.

This definition is not very distinguishing and precise. This is exactly the problem of the definitions of discontinuity that have been proposed (the problem is discussed in more detail in chapter 3). Change can come in several shapes, namely linear and non-linear. Linear change is continuous. Non-linear change may be gradual or sudden change. Non-linear continuous change (gradual) may come as exponential or s-shaped growth.

Exponential, gradual change has no end state. S-shaped growth means that there is an end state. Discontinuous change is either a discrete step or a cusp. A discrete step is an instant change from one equilibrium (horizontal line) to the next. The cusp equals the discrete step, but it has two extras. The equilibria (i.e. temporarily stable periods) overlap, and there is not one, but a collection of possible instant changes from one equilibrium to the next. Sudden change (which is quite often associated with discontinuity) is not a satisfactory aspect of discontinuity.

Stages and sequences: definitions

Instead of stage, words like period, phase, level, point in time, interval and stretch denote a duration of time. These words slightly differ in meaning, but the importance of the concept stage in developmental psychology is enormous. Over time, the ages of man have been a popular concept: man follows a path with or without stages, for instance, from birth to death trough puberty and adolescence. Cicero put it like this: *Cursus est certus aetatis et una via naturae eaque simplex, suaque cuique parti aetatis temestivitas est data* (Life's racecourse is fixed and nature has only a single path and that path is run but once, and to each stage of existence has been allotted its own appropriate quality) (cited in Boom, 1993). In this section, the concept stage is discussed twice: once with regard to the use and meaning of the word in developmental psychology, and then with regard to its theoretical considerations.

The term stage is used (Brainerd, 1978) as:

a. a *metaphor*
b. a *description* (i.e. behaviours that undergo *age change*)
c. an *explanation* of *age-related changes* in behaviour.

Brainerd disputes that Piaget's criteria of a stage are adequate, and although Fischer (1983b) and Fischer and Silvern (1985) mentioned other criteria (synchronicity (i.e. the timing of structure in development), qualitative change, discontinuous change and the presence of limits), there is additional criticism by Glaserfield and Kelly (1982). They hold that *period*, *phase*, *stage* and *level* have been used to name a stretch of time in which, apparently, there are no structural changes. Glaserfield and Kelly (1982) use an orthogonal framework where the x-axis represents a temporal dimension, and the y-axis a quantitative dimension. *Period* is any segment between two values of x. It has a time length, there is a starting and/or end point, and there is an event (state, process) that is characteristic for that period. *Phase* designates a stretch of time, the content of which is in some way related to the content of one or more contiguous segments along the time (x-)axis. Thus, a phase exists when there is more than one stretch of time (as in stage), and some predictable change occurs. *Stage* implies a progression to an end state. Stages are a sequence of segments along the temporal axis, each one of which can be individually characterized by a change relative to the adjacent ones. In this stretch of time something (on the y-axis) is constant, and earlier stages differ from the present one in a dimension other than time. The final concept is *level* which is not a stretch of time, but a specific degree or height of some measurable feature or performance (i.e. measured on the y-axis). In contrast with stage, a level has no time index, only a score (i.e. level is a stage without a time index).

In the theoretical stage debate, Boom (1993) offers a systematic reconstruction of the concept developmental stage. This includes a historical review of the concept. Boom comes up with a characterisation that in a stage both unity and diversity, and stability play a role, and that a stage cannot be a single occurrence. There is a sequence of stages, i.e. there is not **one**, but there are several stages. Boom relates stages to the learning paradox (i.e. the Meno-paradox by Plato), and he advocates a close relationship between novelty in development, stages and transitions.

Although the large body of literature and theories on the concept stage has been discussed, the empirical application of these ideas across several developmental disciplines (including language development) remains problematic. Especially stage transition need more attention. In the debate on stages, Boom (1993) has displayed a philosophical exercise to unravel the problematic concept ‘developmental stage.’ His thesis on the concept developmental stage summarizes the extensive discussion on stage and related topics. However, Boom cannot offer working definitions or even an empirical and mathematical model to describe and explain stages, and the transitions between stages. Furthermore, the concept stage lacks a formal model (e.g. a linguistic model), which leads to the risk of a circular reasoning (e.g. using age related criteria to find stage related criteria).

Preliminary conclusions on the concept of stage

I have adopted the nomenclature by Glaserfield and Kelly (1982). *Stages,* in the meaning of both an end state as well as an intermediary state, is defined as a sequence of segments along the temporal axis, each one of which can be individually characterized by a change relative to the adjacent stages. In this stretch of time something is constant, and earlier stages are different to the present one in a dimension other than time. Furthermore, stages are not age or time related, but they are related to some feature on the y-axis. A philosophical discussion (e.g. Boom, 1993) does not help us out of the muddy pool of the exact theoretical meaning and of the empirical application of the stage concept, it does, however, offer a historical outline to the problem of the concept, that shows the complex matter of the concept over time, and therefore a solution is not easy to give.

2.2 Language: structure and development

Introduction

Theories on general development lack a thorough discussion of the acquisition of language structures (i.e. the acquisition of a grammar). For example, learning theories cannot explain why a child is capable of producing sentences in creative way (i.e. producing sentences he has never heard before). Therefore, we need a model of language to describe and analyse the nature of language structures itself. Linguistic theories offer such a description and analysis. These theories offer a structural analysis, which describes and explains how languages are ordered (on a sentential level), how language are related to each other (i.e. all languages share the same properties at some levels) and why language acquisition is possible at all. Especially generative grammars brought insight into the problems of language development. Grammar descriptions, linguistic theories with respect to surface and underlying structures, and theories on the acquisition of a grammar provide an answer to the question's of why and how language development proceeds in the first three years of life.

The focus in this study is on a subset of development, namely grammatical development. Therefore, findings from linguistic research are emphasised with respect to the *structure* of language. All speakers of a language have an understanding of ill-formed or ungrammatical sentences and words. Since not all combinations of words (i.e. the order of words in sentences) are possible, it needs to be explained why some combinations are not possible whilst others are.

Since children make sentences they have never heard before, a theory has been proposed that is now known as *generative grammar*. This theory is discussed with an emphasis on the structural elements. I surely do not want to go into the details of linguistic theories, since these detailed analyses would lead us away from the main subject (i.e. change).

In the next section, I will summarize findings of linguistic theories[5]. These findings of linguistic (here: generative) theories offer an explanation for the order in language developmen, i.e. a reason why children do not immediately utter adult-like sentences, but go through a series of stages or states. This series of states are highly ordered, but children do not acquire a language by first learning all the words that start with an ‘a’, and then work their way through all other words (from ‘b’ to ‘z’). Additionally, linguistic theories in general and generative theories in particular may explain and/or predict possible shapes of development (in terms of quantitative behaviour). Furthermore, generative grammar is probably the best worked out theory of language, and it relates several languages with each other, by indicating that the structures of various languages are not a random range of structures, but that they are various instantiations of one general underlying level.

Structure in language

Grammars are not new to the 20th century. The Greeks, for example, were fully aware of the fact that language has a structure, which is usually called grammar (from the Greek, **grammatike** > **gramma** for *alphabet* or *letter*). A grammar prescribes language specific rules for a language, that is, it specifies the fixed relations between words within a sentence. For example, the acting person is the subject of a sentence, and the subject agrees in number and person with a verb (in Germanic languages). The grammar of a language can, so to speak, do without any semantic analyses.

Linguistic theories have been inspired by the notion that language is highly organised. Language has a *structure* that has two connotations in linguistics. First, there is the surface structure of a spoken or written language which is the structure given in the language that we hear or read. It is the canonical form of a sentence. This surface structure is ordered according to the rules of a grammar of a specific language. Second, there is a deep structure (i.e. in GB and related theories) which is considered to be the underlying structure. One of the advantages of assuming this deep structure is that ambiguous sentences that have one and the same string (hence, the ambiguity), have an underlying structure that is completely different. A frequently used example might help here.

[5] Generative theories are not the only theories on the structure of language. Levelt (1989) built a model of language processing (not language development), in which he explains how sentences are processed from the initial concept till articulation. Wijnen (1990) showed that errors in language production of children are highly systematic, and they can be explained from other models than generative ones.

Take a look at example 2.1.

2.1. Time flies like an arrow.

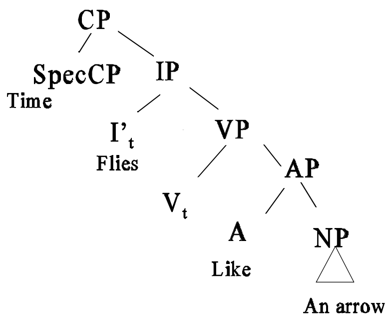


Figure 2.3. Syntactic tree with *Flies* as a verb.

Box 2.1 List of the abbreviations in the trees

CP = Complementizer phrase. CP refers to the complete sentence structure, as a complement of the verb and the objects. SpecCP is the head of a sentence.

IP = Inflectional Phrase. I contains all the features of the verb (like agreement with the subject, both in number and in person).

VP = Verb Phrase. VP is the part of the sentence that contains the objects and the infinite (not inflected) verb. V contains the uninflected verb (e.g. in English, the verb 'be' is infinite in 'to be').

AP = Adverbial Phrase. AP is the phrase that contains adverbs,

NP = Noun Phrase, i.e. a noun and its satellites (e.g. determiners and adjectives).

In recent papers, new phrases like DP (Determiner Phrase) have been proposed, but in the context of this study on change these new phrases would lead us off the track. Therefore, these new phrases are not discussed.

This sentence can be interpreted in more than one way. *Time* can either be a verb, a noun, or an adjective (to *flies*), and *flies* can either be a verb, or a noun. So, although the surface structure is the same, depending on the grammatical interpretation of the words, the sentence has different semantic interpretations (although some meanings are arguably bad: how, for instance, do you 'like' an arrow; poetry often uses this meaning of a sentence). With the help of so called *syntax trees*, it can be understood how the structure of a sentence is organised. In Figure 2.3, the interpretation of the sentence leads to a

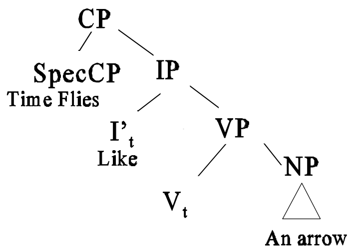


Figure 2.4. Syntactic tree with *Like* as a verb.

tree, in which *flies* is the verb. However, in figure 2.4, the verb is *like*. Using the same basic underlying structure (i.e. the tree), we can explain the two or more different interpretations of a sentence. Note that the syntactic trees are used to elicit these different interpretations. The trees are not explicit and correct with respect to every detail of the syntax of a language, and the present state of theories on syntax.

One of the assumptions of the trees is that the underlying structures in language consists basically of a combination of a **S**ubject, a **V**erb and one or more **O**bjects. Typologically, languages are defined by their canonical word order (e.g. SVO, VSO, etc.). English, for example, is an SVO language. The underlying structure is a basic order of a subject, a verb, and one or more objects. This is reflected in English sentences, both in main clauses and subordinate clauses. However, in Dutch the underlying structure is SOV, which is reflected in subordinate clauses. Main clauses apparently have an SVO, but this order is derived from an SOV order (Koster, 1976).

Let me start with ordering the sentence in example 2.2.

2.2.a The woman saw the man with the binoculars.

There are two underlying structures. A reordering of this sentence with the SVO-idea (the assumed basic order in the English language) reveals that the subject is *The woman* and the verb is *saw*. So S and V are now known. But what is the object? In one case, the man has a binocular (2.2.b), but in the other the woman is seeing the man through a binocular, and we get example (2.2.c):

2.2.b The woman / see (+ past) / the man with the binoculars.
2.2.c The woman / see (+ past) with the binoculars / the man.

The idea of an underlying structure is a crucial element of generative grammars. This idea makes it plausible that language is more than just a string of sounds and sentences. Furthermore, complementary evidence for underlying structures can be found in a discussion on *competence* and *performance*. When adults are asked to, they can perfectly well say which utterances are and which utterances are not grammatical. Apparently, we humans have inbuilt knowledge of language, usually referred to as competence. Even children (at the age of 4) seem to have a knowledge of language when they are offered incorrect sentences that they have to repeat (Ruhland, 1991). In daily life, we do not use all language knowledge we can, for practical reasons. Performance is what we use every day in communication: it is our competence hampered by factors like memory and abilities of the hearer or the reader to understand language.

One of the essential assumptions of linguistic analyses within the generative framework is that apart from words (lexical items) there are functional categories, a special feature of Dutch (and many, but not all, other languages). In a later stage of this study, the relationship between the variables (function words) and the development of syntax is discussed with respect to quantitative changes. Since these functional categories are both important in the structure of a language **and** in this study, I discuss these categories in the next subsection.

Functional categories

The term “functional projections” is the counterpart of the term “lexical categories” (i.e. Nouns, Verbs, Adjectives and Prepositions). Lexical means that they are a part of our (mental) lexicon. Functional categories are projections of these lexical categories. These ‘projections’ are theoretical assumptions of non-concrete linguistic properties (i.e. entities not referring to reality) to explain verb inflections, diminutive and plural forms of nouns and other structural elements of the grammar or syntax of a language. Both categories are phrases in a syntactical tree. For example, it is assumed that the functional category associated with the verb is responsible for a number of features of the Dutch syntax (i.e. Inflectional Phrase). Important is that if functional categories are acquired by a child he has mastered an important step of his language, and his communicative ability will increase subsequently.

Functional categories are associated with syntactic features such as tense, agreement and case. These categories form a ‘shell’ over the lexical categories, which is an interface between the representation that specifies the thematic relations between the verb and its argument noun phrases on the one side, and phonological form on the other. This means that syntactic trees are assumed to invoke functional categories, that most of these categories are invisible, and that they are abstract entities. It means that if these abstract categories are used, expressing thoughts with references to non-concrete entities proves a high command of syntactic relationships. The command (e.g. possible movements of words within a sentence), over these categories is shown in the syntactic trees by, for example, traces (see, for example, V_t in figure 2.4.).

Functional categories are related with function words, and they are discussed in chapter 4. Most important aspect is their syntactic property in a sentence. There are a couple of functional categories (e.g. DP (Determiner) and AuxP (Auxiliary) which are important in this study). The main objective here is to elucidate that these projections are different from lexical projections by their abstract position they hold in a sentence, and their distributional quality in a language.

Provisional conclusions

The previous (sub)sections have been limited to a workable size by first defining the key words of a developmental study. *Development*, for example, has been defined as a change over time. Very useful to the concept of the *stage* is the nomenclature of Glaserfield and Kelly (1982), which is adopted in this study. Stages form a sequence of segments along the temporal axis, each one of which can be individually characterized by a change relative to the adjacent ones. Within this stretch of time the quantity of something measured (e.g. a variable) is constant, and earlier stages are different to the present one in a dimension other than time.

Discontinuity refers to a lack of smoothness in growth curves and of a linkage with earlier behaviour. Although some psychological theories on development assume discontinuity, like for example Piaget's equilibrium model, there are no formal theories or models that are able to demonstrate discontinuity on an empirical level.

Languages have a structure which is commonly referred to as grammar. This implies that language is a rule governed system. It also means that words in sentences have a relationship. Furthermore, in most languages words have a fixed order in a sentence. The assumption of an underlying structure makes it plausible that all natural languages share common features. The surface structure of languages is derived from the underlying basic order through movement procedures. The lexical categories (i.e. words) in a sentence undergo changes (e.g. verbs are inflected). The changes that form to say a shell around the lexical entities are called functional categories. Functional categories are syntactic operations, that do not refer to entities in the real world. It is assumed that these categories are more difficult to acquire.

The importance of the concept of a language structure is that in any language, but also during development this structure follows rules. In other words, language and language development is not without a goal. There is an end state in development, which is described and explained with the aid of linguistic theory. Language development, i.e. the change from no language to that end state, is not a proliferation of change that is adrift, but a series of learning events in time that is rule guided, although these rules do not need to be innate or explicitly learned.

Child language: theories, models and data

Thumbing through the CHILDES/BIB (Higginson & MacWhinney, 1991; supplement 1994), an annotated bibliography of child language (and language disorders), one thing that stands out is the size of the book. It records over 13000 references to research on language development and disorders. Most references show that child language is frequently studied from either a linguistic or a psychological theory. This 'either-or' situation is undesirable, because it makes it more difficult to answer the four core questions of language development research (Weissenborn, Goodluck & Roeper, 1992). These four questions are:

1. What is the adult grammar (i.e. the end state of development)?
2. What does the child bring to the task of language acquisition?
3. What developmental stages are exhibited?
4. What are the conditions for successful learning?

Question 1 relates to linguistic theories (see the **Structure in language** section). Question 2 relates to the problem of innateness (and the learnability problem), question 3 has to do with the developmental problem, and question 4 relates to the continuity and discontinuity discussion. If these four questions represent the core of language development research, which I believe they do, both linguistic *and* psychological theories and methods have to be applied to the study of language development.

The study of language development, and the four questions above, are inspired by the fact that anyone can observe that child language of is different from that of adults (at least from the surface). In fact, children do not speak (i.e. use sounds[6], words and sentences that are the same as adult speech) at all at birth, and their first recognisable word is usually spoken around or after the first birthday. From the surface, child language remains different from adult language for a long time. But if there is a surface

[6] Children babble the same sounds in all languages in their first six months (Van der Stelt 1994).

structure and a deep structure (see previous section), in what respect is child language different from the surface or/and deep structure of adult language, and how do children arrive at an adult level of language production? The assumption of an underlying structure makes it plausible that grammatical development is rule governed, and not an accidental sequence, and that children acquire these rules.

There are two extremes on a continuum of theories, that explain the relationship between possible inborn capacities and the resulting patterns. I start with the *continuity and discontinuity* assumption that is strongly related to what is called *Universal Grammar*. The discussion of these approaches to structural components of language serves two goals. First, it is meant to be a clarification of concepts, and second, both assumptions yield possible different lights on development (in terms of change).

Language development: origin and change

Universal Grammar

The debate on the acquisition of language has been strongly influenced by structural theories. There is also a strong interaction of empirical data with descriptive theories, and with underlying mechanisms of both language and language development. Three questions are important. First, *what* are the linguistic characteristics of language development? Linguistic knowledge must be present in or become available to a child at some moment in time. Second, *how* does a child learn its language? Third, *why* is the order of language development the way it is? These questions are reflected in three kinds of theories:

1. Linguistic theories.
2. Learnability theories.
3. Developmental theories.

Linguistic theories aim at describing what kind of linguistic competence, such as syntax or semantics, a child has. Generative grammar (e.g. Government-and-binding-theory) is an example of a linguistic theory. *Learnability* theories explain how a child is able to learn a language and how a child is capable of learning the grammar and the metalinguistic properties of a language like Dutch or English. *Developmental* theories explain change in language development, that is, the fact that children go through various states towards the adult linguistic state.

A critique on outside-in theories, that explain language development on the basis of cues that a child derives from his language environment, is that the general rules of association, induction and learning cannot explain, for example, the occurence of structural mistakes in (language) development. The following example is taken from Atkinson (1986). Suppose, parents provide a child with the following sentences (examples 2.3 and 2.4):

2.3.a. The man is here.
2.3.b. The man has left.
2.3.c. The tall man has left.
2.3.d. The man who is outside is impatient.

2.4.a. Is the man here?
2.4.b. Has the man left?
2.4.c. Has the tall man left?
2.4.d. Is the man who is outside impatient?

In general the form of the yes-no interrogative is generated by placing the first verb in front of all words (examples 2.4). These examples can be learned from the input without any complex linguistic knowledge. However, in 2.3.d, this would lead to the wrong question (in 2.5.).

2.5. *Is the man who outside is impatient?

Induction from the input cannot explain why mistakes like in 2.5 are not found in the language of children. In the words of Martin Atkinson: "Accordingly, the hypothesis space through which the learner must search in acquiring the syntax of English is restricted so as not to permit hypotheses which embrace rules which are perfectly easy to formulate and completely explicit but which have the wrong sort of formal properties." (Atkinson, 1986, 116).

With respect to inside-out theories, criticism came from Braine (1992, 1994). In his 1994-paper "Is nativism sufficient?", Braine takes up the discussion whether "language acquisition is mostly a realization of innate principles, or mostly a consequence of learning" (p. 10). The problem is that learning needs mental contents or learning principles, which result in new problems. The solution to the problems that both nativist and empiricist approaches of language development are confronted with is *interactionism*, according to which development is the consequence of an interaction between genetic and environmental causes. Braine states that the empirism-nativism argument is a distinction between structure (or mechanism) on the one hand, and idea (or content) on the other. This distinction, however, has been absent from theoretical thinking. Nativism is still the central theory to explain language development, but the problem is that nativism is insufficient because it does not account for developmental change. In the end, Braine proposes that a learning theory based on Piaget's assimilation-accommodation, is needed to explain the ontogenesis of innate cognitive and linguistic primitives. The developmental primitives are:

1. A learning mechanism: old rules analyse new material.
2. Semantic categories.
3. A tendency to classify words and phrases.

Although there is no syntactical innateness assumed, there is an assumption of an innateness of learning principles and semantic categories, and classifying words. MacWhinney, Leinbach, Tabaran and MacDonald (1989) argue that language acquisition is not guided by rules (as in nativist approaches), but that children learn by cues. However, a child first needs to detect a cue (which might be the form or sound, the place of a sound/word in the input), then he must get hold of the frequency of language variables as an index of the importance of the input, the variation of availability, and lastly, he must be aware of the reliability and the consistency of these cues. In sum, the assumption of acquiring language by cues offers identical problems as rules. Since rules offer better descriptions of language development, I would suggest to choose rules, and to use cues in addition when rules fail to explain language acquisition.

Since other theories like Case grammar and Pivot grammar (cf. Ingram, 1989) could not explain the acquisition of syntax either, neither learning theories nor the assumption of full innate knowledge suffice to describe and explain language development (both are too strict to be applied to development). Furthermore, theories on the origin of language development do not explain or even describe the form of change.

Then, the question is: what do we need for a description of language development with respect to universal and inborn capacities? One of the main assumptions is that all languages of the world share so called universals. These shared basic elements of all languages constitute Universal Grammar. What is this Universal Grammar, or UG? How do we get from UG to a specific language? The central issues are the relationship between UG and innateness, and the relationship of UG with respect to development (i.e. in quantitative terms).

Chomsky and his followers have argued that the acquisition of language is impossible without innate linguistic knowledge. This body of knowledge, known as Universal Grammar (UG), is a set of principles (i.e. constraints on the architecture of grammars), that sets borders to the range of possible human languages. Since children may acquire any language depending on their language environment, this leads to a logical problem: how do they do that? Why are all languages learnable for all children? The answer is that it is assumed that languages do not differ in their core grammatical architecture. In this core architecture, the universal principles are parameterised in such a fashion that the designs of specific languages vary over a limited range, each of which implements the underlying principles slightly differently. The child's task is to determine which of the core grammar variants allowed by UG generates the sentences he is exposed to, and to set the values of the parameters accordingly (Chomsky, 1986, Gibson & Wexler, 1994, Goodluck, 1991).

The way UG is embedded in language specific grammars is shown in figure 2.5. All languages are, so to speak, 'surrounded' by UG. Languages differ in their surface grammar, but there is one underlying general knowledge base. In other words, any language is supposed to be a subset of UG. This body of knowledge is abstract and it provides constraints on possible languages. The subset-principle applies: children will always opt for the least permissive grammatical system to handle data in their linguistic environment.

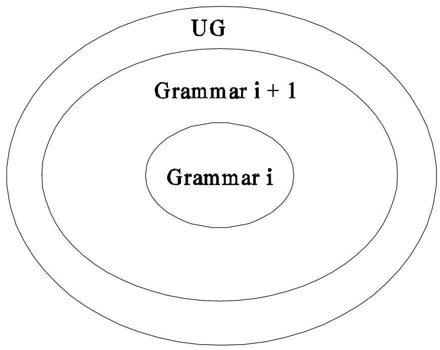


Figure 2.5. Grammar of any language as a special case of UG.

The following quote from Goodluck (1991) expresses the power of the assumption of UG:

> "The use of principles of grammar potentially allows the child to form rules and hypotheses that affect and determine the form of sentence types for which he may have no direct evidence in the input itself." (p. 144/145)

The benefits of the assumption of a universal grammar are that UG enables to predict, first, possible variation in any language (i.e. what sentences children will and will not produce). Second, it leads us out of the problem of learnability. That is, since UG is supposed to be guarding the underlying (however, not specified) structure of language, and it is therefore setting the cognitive and linguistic constraints, the problems of acquiring a language are limited. Third, the presence of UG explains why any child might learn any language. There is one little drawback: UG does not predict the shape of change. That is, although UG sets the boundaries of language development, it does not predict how a score (e.g. on a test, or a frequency count in spontaneous language) changes over age. Is that change linear, or non-linear? Is change continuous or is it discontinuous (in quantitative terms)? Despite UG's explanatory power, these questions cannot be answered by assuming UG.

Change: discontinuity in grammatical development

An important question is whether or not discontinuity (in the linguistic meaning, i.e. a qualitative meaning) predicts non-linear or discontinuous change (in the psychological meaning, i.e. a quantitative meaning). Linguistic discontinuity refers to all theories that assume that some or possibly all principles of UG are not present from the beginning. However, precise predictions of quantitative change, in terms of the change in figures 2.1 and 2.2, are not possible. In other words, linguistic theories are very difficult to link with quantitative analyses, since these theories concern about relationships between words on a non-chronological scale. They do not care about time-dependent issues. The main goal is here to prevent confusion in the rest of this thesis, since the meanings of (dis)continuity differ across disciplines.

A discussion on discontinuity in language development is problematic, since definitions of discontinuity and stage that originate from linguistics are not clear at all. Hyams (1986) notes that an analysis of early language supports the hypothesis of continuous development if it is constrained by principles and parameters of UG. Child grammar differs from adult language with respect to the setting of a particular parameter. In other words, syntactic development is a matter of fixing parameters of UG. Thus, *dis*continuous syntactic development means that each (or some) of the intermediate grammars does not fall within the limits imposed by UG. Discontinuity means that grammars are radically restructured from one stage to the next. These grammars are called semantically based grammars: they lack a syntactic component. The drawback of these grammars is that no simple inductive procedure can explain the transition from semantical grammars to abstract and adult syntax. However, this is in strong contrast with the majority of research, where it is assumed that language development takes place in parallel on all fronts and it does not involve a significant discontinuity or qualitative leap from semantics to syntax. This is also the assumption in this study.

Most of the time discontinuity refers to syntactic properties of language. It means that if grammars are discontinuous they are radically restructured from one stage to the next

(Hyams, 1986). The idea is that each language and every development has (as mentioned in the previous subsection) an underlying structure called Universal Grammar. Discontinuous development means that each of the intermediate grammars do not fall within the limits imposed by UG. Continuous syntactic development is a matter of fixing parameters of UG. A look at figure 2.5 reveals what (dis)continuity is. The grammar of a language is part of UG and if there is no restructuring, there is continuity. The opposite also holds: if a restructuring takes place, then there is discontinuity.

In summary, the difference between several definitions of discontinuity in psychology and those in linguistics is that psychological discontinuity refers to both qualitative and quantitative properties of development, whereas discontinuity in linguistics only refers to a qualitative property. The predictions of linguistic discontinuity are that there must be states that are clearly separated from earlier or later states. This discontinuity hypothesis also predicts all sorts of change. There are no clear arguments why linguistic discontinuity should predict quantitative discontinuity. Only general remarks can be made (e.g. ‘there is an increase over time’).

Change: stage issues in grammatical development

Of the four questions as proposed by Weissenborn, Goodluck and Roeper (1992), question 3 relates to the existence of stages. Let us suppose for a moment that there are no stages (whatever the definition). What is the form of change we might expect in development? In other words, this question deals with the kind of change that may be expected in grammatical development, under the assumption that child grammars are either different from or similar to adult grammar (continuity or discontinuity hypothesis). Theories on language development that may carry a solution for the developmental issue of language acquisition can be found in Verrips (1991) and in Weissenborn et al. (1992). They make a distinction between continuous and discontinuous development (discontinuity is dealt with on the next page). Weissenborn et al. (1992) make a distinction between strong continuity and weak continuity. Weak continuity is when the first child grammars are a subset of adult grammars: there are no changes on the structural level. Fodor (1983) comes close to the strong variant (strong means that all language structures are fully specified from the first grammar onwards), however, the weak version is probably the most popular version. The only claim weak continuity puts forward is that UG principles are obeyed in the grammars during the language development.

Working from the assumption that development is continuous (with respect to UG), the question is what this continuity assumption predicts with respect to quantitative change. Given weak continuity, are we to expect that the quantitative change in the mastery of certain language variables takes the form of a continuous curve? The continuity hypothesis (i.e. in linguistic terms) does not a priori reject discontinuous change (i.e. in quantitative terms). If we suppose that continuity is the right hypothesis for language development, then this approach allows for discontinuous change, since continuity does not rule out the presence of stages. The only claim continuous models make is that the assumptions of UG are not violated. The prediction on discontinuous quantitative change is not narrowed by the continuity hypothesis.

It is important to discuss linguistic hypotheses to elicit the expected quantitative predictions on development. These hypotheses and their prediction on change demand a careful treatment to prevent problematic interpretations of the literature in the rest of this thesis. Two extremes have been discussed (continuity and discontinuity), but there is a third issue, namely the existence of linguistic defined stages, and the transitions from one stage to the other.

From stage to stage: transitions

Since language development has its own characteristics, there are with respect to the meaning of the concept stage differences between a general psychological and a linguistic approach to stages.

Ingram (1989) dedicates chapter 3 of his book to stages in language development. He points out that has *stage* rarely received a thorough definition. He makes a distinction between single and multiple behaviours. Single behaviours refer to the situation in which one behavioural variable is measured over time. In this case there are four kinds of stages. According to Ingram, a *continuous* stage is a stage ‘where a single dimension of

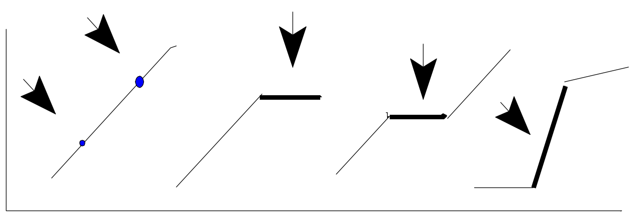


Figure 2.6.a b c d. From left to right: continuous (2 dots), plateau, transition and acceleration stage

behavior is being observed' (p. 33). Graphically this looks like figure 2.6.a: it is a point on a continuum. Since I applied the definitions of Glaserfield and Kelly, this is not a definition of a stage, but a definition of a level. The second stage Ingram distinguishes is a *plateau* stage, and it is the end stage of any language behaviour. In other words, the continuum is halted. The plateau stage is depicted in figure 2.6.b. The third stage is called a *transition* stage: behaviour is expected to change again at some later time. However, I think this is a strange definition. It is a stage in terms of Kelly and Glaserfield, but whereas most transition definitions refer to sudden change, Ingram refers to stretch of time where there is no change at all (see figure 2.6.c). It is an intermediate plateau stage. The *acceleration* stage in figure 2.6.d is a rapid change between a 'transition' stage and a 'plateau' stage. There is a problem: if an acceleration stage looks as steep as a transition stage the difference between these stages is impossible to tell[7].

At the end of the chapter, Ingram bears to mind the distinction made by Brainerd (1978). Like Brainerd, Ingram distinguishes two kinds of stages (apart from the four specific stage definitions on the previous page). Descriptive stages are stages with behaviours that undergo change and there are antecedent variables that are responsible for the change. In explanatory stages the variables must be subject to independent measurement.

In sum, the stage concepts that have been proposed by Ingram are confusing: there is a contradiction with the use of transition stage as being a stage of no change. There is also the problem of applying these stage definitions to an empirical study of language development, since these definitions of stages are qualitative, and a developmental perspective is missing. Furthermore, the definitions remain non-mathematical (and therefore non-testable in the sense of non-linear theories). Linguistic stage definitions have been discussed so that the study of language development may benefit from it. Furthermore, since this thesis is on the verge of linguistics and psychology, the concept stage (especially the differences between the definitions in both disciplines) was discussed to prevent confusion.

Since it is assumed that there are some stages (e.g. the prelinguistic and the linguistic stage), the question is how a change from one stage to the next takes place. Transitions (i.e. a change from one stage to the next) have been given different meanings and explanations. For example, *transition* in language development has been used for the language of a child that changes from the non-linguistic to the linguistic stage. Dore, Franklin, Miller and Ramer (1975) distinguished two transitional phases. The first is in

[7] Ingram also discusses the issue of two or more principles that cause behaviour (multiple stages; Ingram, 1989, 37). However, this issue is not relevant to this study.

agreement with Camaioni (1989), Golinkoff (1983) and Sugarman (1983) who state that there is a transition between prelinguistic vocalization and one word speech. None of the authors clearly defines this use of the word *transition*. The second transition would be between one word speech and patterned speech.

Dore et al. (1975) postulate a claim regarding transitions in language development. Each major linguistic stage, they say, is preceded by a transitional phase which serves as a bridging device for the next major acquisition. Furthermore, transitions are extremely important, because their delineation is essential to a complete description of language development and they relate directly to basic theoretical issues concerning developmental progression. According to White (1982), a transition is a matter of changing perception. That is, a change in the language (called a transition) is caused by a change in the way a child is perceiving the parents. If the parents change either slowly or radically, the effect on the child should be immediately evident. Whether this is the case should be evidenced by empirical data.

The problem is that transitions, for example from non-linguistic to linguistic behaviour, are loosely defined. In fact, transitions in language development are changes between two stages. This change is described in terms of a *qualitative* change. Although it is assumed that this linguistic transition is abrupt, there is no quantitative fundament of this abrupt change, and the discussion on this concept of transition is build on disputable arguments, since it lacks empirical data to decide for or against abruptness. In this study on grammatical development, I will come up with a definition.

The relationship to linguistic knowledge is that the development of linguistic structures (e.g. functional categories) in terms of branches of the linguistic tree depends on the assumption of the underlying ‘knowledge’ or syntactic structures. If all knowledge is present (the overall syntactic structure is available to the child), development is presumably a gradual or even linear process. The other extreme is that no syntactic structure is available, and the prediction of quantitative change is a sudden development, since syntactic branches do not ‘grow’ in a gradual way. In other words, branches are there or they are not. In chapter 4, this discussion of this underlying syntactic competence is taken up, and related to the variables in this study.

Provisional conclusions

In sum, there is a difference between the psychological and the linguistic meaning of stage and transition. *Stages* in child language, as discussed by Ingram, show a resemblance with the definitions of Glaserfield and Kelley (1982), but Ingram reserves the transition stage for what I call an intermediary stage or end stage, whereas an acceleration stage is used for what I call a transition stage. In linguistics, *transitional* change is reserved for the change from prelinguistic vocalization to one word speech. A second transition would be between one word speech and patterned speech. As I pointed out, none of the authors clearly define the use of the word *transition*. One of my aims is to come up with a definition of transition in (language) development (see also chapter 3 and 6). The second goal is to define transitions and stages so that they are empirically testable. Although the definitions in the literature are clear, they cannot be applied to language development to find evidence for (dis)continuous change.

The assumptions of linguistic continuity and discontinuity, the idea of underlying knowledge called Universal Grammar, and questions on the absence or presence of stages do not yield any a priori prediction of the psychological, i.e. quantitative, (dis)continuity. The choice for some sort of underlying grammatical knowledge of a child are therefore not determinative for the sort of quantitative change one might expect in development. Therefore, linguistic analyses of development and of the question of innate structures must be chosen on grounds of structural analyses, not on grounds of expected quantitative change. The reverse reasoning also applies: the choice for the assumed underlying knowledge of a child does not depend on the quantitative patterns that are to be found.

2.3 Summary and conclusions of the chapter

When I gave this chapter the title *Everything must go*, I had in mind that a child develops his language simply because he has to. A child acquires a language, maybe because there is some 'software program' in the brains, probably because language is easily acquired (i.e. without explicit teaching) and it plays a crucial role in the communication of humans. In either case, development will go, and it has to go. Both linguistic and psychological theories aspire to explain why language is easily learned, but that is not enough to fully comprehend language development. What is needed is a model of the *change* of language *structures* in the course of *development*. That is, we need an account for the change of every aspect of a language and why these aspects 'must go', i.e why nothing remains the same in language development.

I discussed (language) development from very general observations and theories all the way down to a review of linguistic theories and findings. In section 2.1, it turned out that theories about psychological development have been focussing on two aspects. First, when describing (psychological) development, there is an assumption of *change*. Developmental theories either seem to support *stage wise change* (development), or *development without stages*. This dichotomy has serious implications. Non-stage theories account for change simply as a linear dependency of time and input, i.e. an additive feature since there is no reorganisation of (underlying) structures, stage theories sometimes implicitly assume a reorganisation. If there are stages, the mechanisms for a stage shift are crucial. Second, the origin of development has been one of the central topics of development research, and it is a dichotomy, too, called the *nativism-empirism debate*. In short, this debate has to do with how much we humans are endowed with when we are born. Nativist theories assume that somewhere in the brain humans have inborn capacities. Empiricist theories reject the presence of any inborn capacities. Finally, the discussion on (dis)continuity has not had clear-cut proposals for the field of developmental psychology. The general overlap between various definitions is that *discontinuity refers to a qualitative change*, i.e. there is no connection with previous moments in time with respect to behaviour. If development is discontinuous, there must be stages in development. However, the discussion on stages is complicated by a lack of clarity in nomenclature. The nomenclature by Glaserfield and Kelly (1982) is adopted in this study. A *stage* is a segment along the temporal axis. Each stage can be individually characterized by a change relative to the adjacent ones. In this stretch of time something (on the y-axis) is constant, and earlier stages are different to the present one in a dimension other than time. Third, the question on (the mechanisms of) *transitions* is directed to an explanation of how and why a system is forced to a next stage.

General psychological theories are not suited for a specific treatment of language development. The study of language behaviour, for example in the behaviouristic tradition, is not the study of language *structure*. Learning theories cannot explain the complexity of language. The idea of underlying structures of language, relations among languages and others has become known as generative grammar. Syntax (i.e. the structure in a language) is a highly organised system, and relations between words in sentences can be described terms of binding and governing relations. In fact, language is a system ruled by parameters and principles. Syntactical analyses within a generative framework offer justifications for the *order* in language development. However, the *shape* of change is hardly predicted by these theories.

Subsequently, I discussed theories on language development. Especially the continuity-discontinuity debate in linguistic theories constitutes an enormous quantity of literature. *Continuity* means that there are no grammars of a child that fall outside the limits that are set by Universal Grammar (UG). UG is the starting point for all children. It is a body of general, non-language-specific principles, from which all languages are derived. This view is adopted in this study. *Discontinuity* in linguistic theories is used for those grammars that do not build upon the principles of UG (therefore, the meaning differs from the one in psychology).

The linguistic s*tage* definitions, given by Ingram (1989), are not unequivocal. They are in contrast to the psychological definitions. Linguistic stage definitions are qualitative, a developmental perspective is missing, and most important, the definitions remain non-mathematical and therefore hard to apply to development. *Transitions* have been allocated to two moments in language development: the transition from prelinguistic to linguistic speech, and the transition from one to multi word speech. Both stages and transitions are not well defined, and mechanisms of transitions are even less well defined. However, both stages and transitions are assumed to be an important aspect of a theory of language development.

The reformulation of the research questions of the first chapter means that core problems are reanalysed. Concluding the present review of the literature, the question becomes:

> If some theories on human development propose discontinuity and others assume continuity as the quantitative pattern in development, and under the assumption that underlying properties of UG and principles are either available to a child from birth on or they become available during development, what are the sorts of change in language development in terms of a relationship between a change on the time-axis and a change on the score-axis?

In the next chapter, I discuss ‘Non-linear theories and models’. A specific class of these models is called Catastrophe theory, that can be used to demonstrate discontinuities (i.e. in the psychological meaning) in (language) development.

3

Ard: Non-linear models and theories

Ard, guth is saorsa	*High, a voice and self-determination*
Ard, tha sinn gluasad	*High, we are moving*
Gu h-ard theid sinn suass	*We will reach up there*

Ard (High) by **Runrig** from the CD **Amazing things**

3.0 Introduction

High, Ard in Scots-Gaelic, is a place somewhere above. For a child to get high, i.e. to grasp grammatical knowledge, he has to move, he has to change the low into a high. Standard psychological and statistical methods, and methods for the analysis of empirical data assume a linear underlying form of change (e.g. linear regression) and a unimodal distribution of scores. Suppose that at least some parts of (language) development are basically non-linear or even discontinuous, and stages (i.e. a multimodal distribution in empirical data) exist, what models can be used to prove that the development of child language is a non-linear, or even discontinuous process? This question is answered in this chapter.

The crucial discussion in the previous chapter is the one on discontinuity and continuity. This discussion has not been settled yet, due to the absence of criteria and a meaningful definition of discontinuity. In chapter 2, I offered a working definition of discontinuity, which is:

> There is a qualitative (structural) change in development in such a manner that there is no structural linkage (i.e. after the change there is a structure that was not there before) with earlier behaviour, and this change is reflected by the lack of smoothness of a growth curve.

This definition is not very useful if applied to language development. For example, what does, in empirical terms, ‘no structural linkage’ and ‘a lack of smoothness’ mean? Thus, a discussion on discontinuity is more or less futile because of a lack of a conclusive definition and method. However, a new research method has been introduced in (developmental) psychology in roughly the last twenty years. This method aims at providing criteria for the study of discontinuities, and it embraces the concept of non-linearity. Non-linear models proved to be fruitful in (developmental) psychology, and they are discussed in this chapter.

In section 3.1, continuous linear and non-linear models and theories are discussed with a special interest in the application to (developmental) psychology. These models and theories are used to solve developmental problems. I shall explain the non-linear equations of these models.

Some theories of human development, for instance Piagetian theory, assume stage wise (or discontinuous) development. However, a model and theory of such development with unambiguous criteria has been absent. Catastrophe Theory has been developed in order to demonstrate discontinuities in development (section 3.2).

In the summary of this chapter, the basic question of how to explain change, in particular non-linear change, is reformulated, with the stress on quantitative aspects (instead of qualitative, i.e. linguistic, ones).

3.1 Continuity: linear and non-linear models

Introduction

Non-linear models began to flourish when Lorenz and his collaborators ran a series of experiments. These experiments were carried out in the sixties, and they involved extensive calculations of weather forecasts. Due to some small accidental changes in initial values, Lorenz found enormous changes in the end values. He and his collaborators ran these calculations again because they thought that they had made a mistake somewhere. But they had not. Their calculations lead to the discovery that small changes in the initial state lead to big differences in the end and that weather forecasts only have a limited predictive value. This discovery, ‘small changes may lead to big differences’, is the ‘butterfly of Lorenz’: when a butterfly in Brazil strikes its wings, it may eventually cause a thunderstorm in Tokyo[1].

Lorenz’s finding has indeed caused some strong winds. However, they did not take place in Tokyo or any other particular place, but they affected a host of scientific research programs (e.g. patterns in fluctuations on the capital market can be explained by using non-linear paradigms; cf. Peters, 1991). In psychology, the application of catastrophe (from the Greek **kata** for ‘down’ and **strophé** for ‘turn’) theory by Zeeman (1976) has been propitious for the study of sudden changes in the behaviour of animals (i.e. dogs). In developmental psychology, Thelen (1989) and van Geert (1991) are strong

[1] An introduction into chaos and non-linear processes is given by Gleick (1987) and Stewart (1989).

advocates of applying dynamic systems models to the development of language and to motor development. The application of these models was also inspired by the fact that traditional statistical methods did not suffice (i.e. the failure of the linear paradigm; Peters, 1991).

Two important issues are discussed in order to be able to apply these models to language development. First, I discuss dynamic systems models (DSM) and dynamic systems theory (DST) that have become popular in a very short time. Second, dynamic system models are built up from difference or differential *equations*. I discuss the elements of the equations that constitute a mathematical model of cognitive development (van Geert, 1991), and the results of existing modelling.

The reason to use dynamic systems models is that they can be used to describe behaviour in biological and chemical processes that show chaotic behaviour, that is, behaviour that does not follow smooth paths along the time continuum. By using dynamic systems models in the form of differential equations, scientists were able to explain the behaviour of biological dynamic interactions like predator-prey-models. In fact, all behaviour which has a time component and all behaviour which cannot be explained by standard linear theories may profit from a non-linear approach.

Dynamic systems models are not chosen on arbitrary grounds, but because they are convenient in four ways. First, empirical data can be fitted to decide what sort of change one is dealing with. That is, a fit with a growth model could help explain the shape of development. *Fits of developmental curves* can be a first help in deciding wether or not change is non-linear. Second, these fits have to be explained. That is, why is non-linear behaviour present in development and especially, what are the *mechanisms* responsible for this sort of change? Third, with the aid of non-linear equations and theoretical considerations a *model* can be made to describe development. But this description is more than a description if the model is correct, because, fourth, *predictions* can be made on the basis of the model. So, these models have descriptive, predictive and explanatory adequacy.

Analyses of development are based on theories or on quantitative behaviour. This quantitative behaviour in development (closely connected to the growth curves of Sternberg & Okagaki, 1982 and to the analyses of Fischer, 1983a) has a number of characteristics. First, variables change over time. Irrespective of their increase or decrease (or even stay on the same level), there is a time index (in graphs, this is usually the x-axis). Second, the variables have a score over time. That is, a variable can be measured (the score is on the y-axis). Third, from these quantifications, qualitative derivations can be made to discuss whether discontinuity is present in development. The crucial aspect is that models are *quantified*.

The definition of development I use is “a score of a variable changing over time”. At any moment in time, there is a score (e.g. the outcome of an observation, or a test score). This score either increases, decreases or remains constant. These scores result in a growth curve (i.e. a collection of points connected with lines). Traditionally, this collection of points (or the resulting growth curve) is explained and described by linear models.

Linear models

In psychology, linear models have been used to explain the underlying trend in a series of data. Most important aspect of linear models is that they are continuous and that the independent variable is additive. Here, I confine myself to models of change that are time dependent. Thus, the time effect of the next point is added to the present point. In other words, there is linear relationship between the independent (time) and dependent variable(s).

The classical linear regression model, i.e. an equation for a linear relationship between the dependent and independent variable, is in (3.1).

(3.1) $L_t = a + b \,.\, t$

An additive effect means that the effect of time on the course of a dependent variable is at any time equally high.

(3.2) $L_{t+n} = a + b\,(t + n)$

The equation means, in words, that the level of a dependent variable is the sum of a constant *a*, and a term *b* (the growth ratio) times any time index *t* (e.g. age). The equation is additive in the sense that for every value of *n*, the time effect is additive. The dependent variable is dependent on time, not on its own previous value.

Linear equations do not necessarily take the form of a line, since polynomial equations are also linear. The generalised quadratic linear model, which allows fits of curvilinear changes, is equation 3.3. Like in equation 3.2, equation 3.4 shows a linear effect for every *n*.

(3.3) $L_t = a + b \,.\, t + c \,.\, t^2$

(3.4) $L_{t+n} = a + b \,.\, (t + n) + c \,.\, (t + n)^2$

New terms (e.g. d . $(t + n)^3$) may be added to the equation, but the relation is still linear, although it is a more complex polynomial equation, in which time is the factor where development is depending on. In the last five years or so, additional continuous models have been developed to explain development. These new models are non-linear.

Non-linear, continuous models

A dynamic systems model of cognitive development

A number of studies assume that non-linear trends exist in development. These trends have been called non-linear because the trends of growth curves are, with respect to their growth over time, either gradual but not linear (cf. Dromi, 1986 on the development of vocabulary) or they occur suddenly (cf. Brown, 1973 on the development of function words). Non-linearity means that age or time in general cannot be the independent or explanatory factor. Factors that determine non-linearity are treated in detail in this section.

The non-linear theory that is discussed as an example is a growth theory (van Geert, 1991). The growth model proved to be an adequate model of development. With the aid of an example from the literature (Wijnen, 1996), I discuss the conceptual elements of the model (i.e. growth rate, resources, carrying capacity), their relationship with psycholinguistic entities, and a mathematical model of behaviour.

It is important to notice that any model is a simplification of the real human behaviour. In this chapter, this simplified behaviour (i.e. a model) is, however, not just all behaviour of all children (for example, averaged over age). The model is build on the basis of *individual paths of development*. Furthermore, any model can be based either on theory or on data or on both. Only a theory or data is presumably not enough to fully comprehend the nature of development. Using data only could lead to a nice fitting result, but any highly complex model could be used to fit data perfectly. A model of non-linear equations that is based on a theory only might have fitting results that do not follow the path of development.

An extensive treatment of how to apply dynamic systems modelling to cognitive and language development can be found in van Geert (1991). To explain the parameters in the model and also how to build a model of (language) development, I use the following example throughout this section. Suppose that a child is learning a rule (in this case a rule about language), namely the rule ‘The inflection of verbs. The total amount of knowledge (i.e. the knowledge of all situations where this rule applies) can be expressed as a 100 % score on a test. It is debatable whether adults will use this knowledge in all situations correctly, but suppose this is approximately the case. The process of arriving at this 100 % score, the change, might be linear or gradual, or even a sudden jump. The mathematical equations to model this process of inflection consist of several parameters: a starting value, a growth coefficient, a final state, and (possibly) supportive and competitive factors (since it is assumed that processes are not independent from sources or other variables or processes).

One of the main concerns of this study is how to account for quantitative growth in development and how to relate this quantitative growth to qualitative analyses and to the parameters in the model. The definition of such quantitative growth (or development) is "an autocatalytic quantitative increase in a growth variable following the emergence of a specific structural possibility in the cognitive system." (van Geert, 1991, 2). In other words, language development (or cognitive growth in general) consists of growth which is expressed in numbers (quantitative) and which is not entirely caused by some external factor (autocatalytic). This growth is caused by the system itself, i.e. change is induced by the cognitive system itself. In the case of finite verbs, the growth of finite verbs is, for instance, not induced by the development of verbs in the input (the parents), but it is due to the structure of the system (i.e. language) itself.

There are a number of assumptions in van Geert's mathematical model. First, it is assumed that language growth is constrained by limited resources. Memory, attention and motivation (among others) are limited during development. To use the example of verb inflection, the speed of learning to use finite verbs is restricted due to these resources. Second, there is an end state of the language system (in the context of this thesis prescribed by a syntactic theory) which is a collection of supportive and conditional factors. These factors result in the carrying capacity, which is in our case, the knowledge of the application of verb inflection that can be achieved by a language learner who is subject to certain limitations of learning and acquisition. Third, growth is hampered by what is called feedback delay. Feedback delay is the effect of a specific growth level on a later growth state. If a child learns to inflect certain verbs, for example, in Dutch so called regular (or 'strong') verbs that are inflected with -te or -de. An example is in 3.1.a.

3.1.a Oom Willem bakte zoekte broodjes.
'Uncle Bill baked little sweet bread.'

A child that has learned the inflection rule (in the example: *bak* + *te*) will benefit from this knowledge on later growth states. It may even lead to regular inflected irregular verbs like in (imaginary) example 3.1.b.

3.1.b *Tante Hillary loopte de keuken uit. (* means ungrammatical)
*'Aunt Hillary walked out of the kitchen.'

Fourth, there are supportive, competitive and/or conditional relationships between growers (since language variables use the same limited resources). Supportive relationships are, for example, the inflection rules for nouns, competition may be found in the two forms of verbs (regular and irregular; see also example 3.2) and a condition for the inflection of verbs is the recognition of verbs is being verbs. A mathematical model of cognitive growth has been built, using empirical data of Dromi (1986) (van Geert, 1991).

In the next section I shall explain the transformation from theoretical or empirical considerations to a mathematical model, based on equations.

Non-linear equations

One of the key concepts of growth models is the end state. The use of a rule (e.g. inflection of the verb) is at the end state, ideally, a 100 % score. This means that the full rule application is an equilibrium. The level parameter x_t, which is in our case the relative number of correct inflected verbs at time *t*, is growing from one equilibrium (the initial state) to the carrying capacity (the end state). It is also a time dependent variable (the index t), that is, all growers in a growth model are time series variables (e.g. age). I assume that development is an *iterative process*. That is, present behaviour is depending on previous behaviour. The process depends also on a *growth rate r*. That is, the speed of change of the process is depending on a parameter r which influences the variable x_t. In the linear equation in 3.1 this parameter is b. The growth rate may be a fixed integer (which is individually defined, and therefore not depending on other factors), but this growth rate may also be a composition of several factors, like memory or other growers. The last option is more realistic, since we assume that no process stands on its own. The various forms of change (linear, gradual or sudden) are expressed by a *power parameter p*. In van Geert (1995), it is explained that the power parameter p specifies to what extent L affects the further growth of L. If p = 0, equation 3.5 describes an restrictive process, whereas for p = 1 a logistic and for p = 2 a sudden process is described. Since p may be any real number bigger than zero, an unlimited variety of forms can be described by the equation.

The equation so far will lead to exponential growth (with r > 1 and p > 1). Of course, this is not in line with normal development which is limited by constraints (e.g. memory). There is an end state in language development. This end state is the body of grammatical knowledge, and it can be built into the equation. In the model, the end state is called the *carrying capacity, K*. The capacity expresses the more or less stable situation: it is not a constant equilibrium, since it might fluctuate due to random influences. The set of principles discussed so far can be transformed into the following equation:

(3.5) $L_{t+1} = L_t + L_t^p * (r - r * L_t / K)$

for $p \geq 0$, $t = 0, 1, 2, ..., n$, and $L_{t=0} > 0$.

One of the strong advantages of these models is that relevant psychological entities are embodied in these equations[2], for example, the dependence of behaviour on its previous state(s) and the dependence on external or internal resources. But there is also a possibility of adding competition and support to these models. This competition is any other grower that influences the grower of study. Competition is when two variables in the same system (language) use the same (e.g. psychological) resources in the same temporal space. Suppose that the plural morphology of nouns calls on the same memory units as the inflection of verb. If all attention is used for inflecting verbs, nouns have to 'wait' before they get pluralized. This competition can also be found in problems of sentence planning. It seems that when children make longer and more complex sentences, they start to have difficulties in planning the sentence. The problems of planning difficulties result in developmental stuttering (Wijnen, 1990). The competition between attention, memory, etc. on the one hand, and the execution of linguistic rules on the other hand may be added to the model as a factor C_t. Models of competition and support between several growers are called models of *connected growers*. An example of connected growers is the development of multi word sentences. If a child has no words, he cannot make sentences. Words are building blocks of sentences. Thus the development of sentences is connected with the growth of vocabulary. This modelling of connected growers, i.e. a sequence of stages, on the basis of dynamical systems theory has been applied to solving cognitive problems (Fischer & Granott, 1995) and to the development of knowledge of children (Case, 1992).

An important question is whether K, r, p, and x are changing over time. Should it be assumed that all parameter values are constant over time, or does, for example, the growth rate r change during development, due to the increase of memory capacity? It seems that the latter is more likely (because of an increase in memory capacities, or

[2] There are two sorts of non-linear equations. *Differential* equations are used when the limit of the time interval is zero, i.e. there are no intervals between two points of time. *Difference* equations are used when there is a discrete time interval. Research on development is discrete by definition, since there is no possibility of measuring every moment, although the processes may be continuous (unless they are the result of discrete learning events).

teaching and learning). An increase of K (i.e. a new equilibrium) implies that the system matures. For example, the child has somehow deduced the rule for verb inflection. The knowledge of this rule leads to a new equilibrium. In a model of development, it is possible to use either constant or variable parameter values. The choice depends on theoretical or empirical considerations.

Everything that grows needs a ‘seed’. In a growth model of language development this seed, needed for the language system to start functioning, is the smallest unit in a collection of sounds, syllables, words, sentences, etc. In our model (Ruhland, Wijnen & van Geert, 1995), words are assumed to be the smallest unit. The seed in our model is a word, which is represented by an equation containing a small number slightly bigger than 0.

Various assumptions of feedback delay are needed since actions in the past may not have direct consequences for the presence. This “credit card effect” called feedback delay also needs to be build in, too, since this delay keeps the system from increasing exponentially (in accordance with empirical evidence; see for example figure 4.1 in the next chapter). Feedback delay is defined as the time interval between a growth level and its cause (van Geert, 1991).

Sometimes, a child undergoes a temporary decrease in his performance. This is called *regression*. The score of a variable decreases and the system quite often loses its stability (i.e. the system becomes *temporarily* unstable). However, regression does not need to be implemented in the equation as a separate, explicit parameter. Consider a system of two connected growers (e.g. two opposing grammatical rules: the use of verb inflection vs. the use of pronouns). When the new grower starts to grow, this new grower (or rule) might lead to a temporary drop in the performance of the first rule, or lead to a complete disappearance of the first grower. For example, when a child learns a new rule (e.g. the inflection of verbs), the child has less time (in terms of memory or attention) to spend on, for example, lengthening his utterances.

Finally, two concepts are used to describe how patterns arise in a system. First, there is *selforganisation*. Selforganisation is when a pattern arises in a system, but when no specific rules are given that prescribe the pattern (see Kaufman, 1993 for some examples). Second, an *attractor* is a state of a system where it is forced to without any explicit instructions. In physical terms, the system searches for its energy minimum. In psychological terms, this means that there are states in a system which are easier to reach, either because of the input, or because of the constraints of the system. Selforganisation and attractors may also be the key to get out of the logical problem of language acquisition. These attractors may be, for instance, based on competition (MacWhinney, 1997).

Applications of the dynamic modelling have been applied to cognitive development (van Geert, 1994a, 1995). Similar and equally promising results have been found in other lines of developmental research. Thelen and Ulrich (1991) found evidence for underlying non-linear processes in motor development. Children have walking skills in the first months of life. This behaviour is present from birth, but it may be dimmed by other factors (e.g. too much fat in the child's legs, or not enough muscle power). In McCune (1992), the dynamic systems view is used for the development of first words, while Fogel (1992) used the dynamic systems approach for the dynamics of movement and communication. In sum, the dynamic systems approach is useful for all research which aims at describing and explaining processes that do not follow a smooth path along the time continuum.

Non-linear equations form a new collection of models of psychological processes. The purpose of these equations is not just a theoretical exercise, but they are used to fit data and build models. The goal of fitting is to find evidence for a certain kind of change, and to give an interpretation of the equations, and the processes they fit on to.

Dynamic systems equations have three applications. First, there is the possibility of *curve fitting*. Fitting data means that parameters of equations are estimated to determine the basic form of the development of a grower. In this sense, non-linear models have a statistical purpose. Most of the time, data are tested against assumptions of linear regression and unimodal distribution (see the next section on other than unimodal distributions), but data can be tested against non-linear regression models.

Second, empirical curves can be *predicted*. There are two methods for predicting empirical data. So-called *genetic algorithms* take as an input a few data points and calculate the next (unknown) data points[3]. This method is extremely useful for, for instance, prognoses on the capital market. Second, predictions may be based on *connections between growers*. This means that if the relation between two growers is known (i.e. extracted from theoretical considerations), this relationship can be modelled with non-linear equations. Running simulations with these connected growers shows how a change in the development of one grower leads to change in the other. For example, if for theoretical or logical reasons multi word sentences (sentences longer than one word) can only arise if vocabulary has developed satisfactorily (i.e. vocabulary is over some threshold), then the grower "vocabulary" is a precursor of multi word sentences. The precursor of vocabulary, for example, is dependent on phonological and phonetic structures and processes. In sum, every variable has a precursor that is a

3

Neuralyst is a program that uses genetic algorithms. However, it would lead to far to go in full detail in the theory of genetic algorithms in the context of this book.

summary and reduction of other systems or partial subsystems (e.g. memory, neurons, attention, etc.). In the predictions of the change of growers, this reduction is necessary to avoid an infinite regression of precursors (with respect to the precursor of vocabulary, see McCune (1992) for a study on the dynamics of first words).

Third, equations are used to build *models* of development. These models can explain how certain behaviours arise, but this explanation definitely needs a structural theory which specifies possible structures. In language development, these possibilities are given by a linguistic theory (e.g. by trees). For example, the initial sentence structure (CP) of a child might be built up from an Noun Phrase (NP) and a Verb Phrase (VP) (e.g. a subject and a non-inflected verb). If we take as a start a reduced competence hypothesis (see for example Radford, 1988b; see also chapter 4) which entails that some branches in the linguistic tree are not available from birth (only lexical categories like NP and VP are available), the question is in what fashion new branches are added to the tree. Do they grow instantaneously and all at once? Some hypotheses assume that all functional branches of a linguistic tree are available from birth (a linguistic 'strong continuity point' of view), but whatever the assumption, development itself might be hampered or limited by memory or attention. The growth model and the linguistic theory are connected as two complementary issues. A linguistic theory provides a valid argument for the choice of the variables, and the theory also specifies the relationships between these variables. In other words, a linguistic theory specifies the structural terms that are translated into quantifications in a growth model. These linguistic theories also explain to what degree the assumptions of innateness predict the course of development. To recapitulate, a growth model must take into account a structural theory that explains the order of development. The structural theory is also needed as a source for the choice of the variables.

Summary of continuous non-linear models

In sum, growth models are complementary to structural models, i.e. linguistic models. In growth models, a process is split up in various elements that are based on theories of development. Through the assumptions of these theories these elements are translated into parameters that are linked to a psychological reality. These parameters in non-linear equations (the important element in growth models) interact with each other, and describe quantified behaviour (in contrast with structural models that are qualitative in nature). Dynamic systems theory and growth models explain the interaction between a grower and its resources, the connection(s) between growers, and the characteristics of a process. Linguistic models and theories are needed to choose the variables that express the development of language structures, and these models explain the order, i.e. the way words, constituents (e.g. 'in the pub' in the sentence 'I am drinking a couple of pints in the pub') and sentences are acquired by a child, of the acquisition of language structures in development.

The models that I presented are continuous, but non-linear models. These models exhibit strong accelerations, regressions, and fluctuations. Growth models are also very useful for modelling relationships between several variables. However, these models are continuous models, so discontinuous behaviour cannot be modelled. That is the reason to introduce Catastrophe Theory in the next section

3.2 Evidence for discontinuities in processes: Catastrophe theory

Introduction: What is Catastrophe Theory?

Brown (1973) and others found that function words (see chapter 4 for an extensive treatment of these words) appear suddenly in the language of children. The problem is how to define *suddenly*. The concept of suddenness in the development of, for instance, function words remains vague, since it often entails the notion of discontinuity. To overcome the vagueness, discontinuous models are needed. The models discussed in section 3.1 can be applied to processes that are non-linear. Van Geert (1991) has elaborately shown that his models are adequate for the description of cognitive development (e.g. language development). However, these models cannot be used for the identification of transitions (sudden changes) between states: these models cannot discriminate between accelerations and true discontinuities. A subset of non-linear models that can be of help for the explanation and description of transitions (sudden changes) is a theory that offers criteria for transitions and states. This theory is called Catastrophe Theory (CT) (Thom, 1975). CT is a mathematical theory of discontinuous equilibrium changes (or phase transitions). This theory provides empirical criteria (called flags) for the detection of phase transitions. It also provides a formal definition of stage wise development. Since the stage criteria of Piaget have met with substantial criticism, CT can be used instead in order to provide a formal model of developmental transitions and stages. This theory offers criteria that have to be linked with empirical criteria. Note that the definition of discontinuity is the psychological definition, not the linguistic one (see chapter 2). The graph belonging to this form of discontinuity is Figure 2.2 (the right figure).

In short, CT is a theory that explains the characteristics of a phase transition in a system. During a transition, the system is unstable and therefore the system shows features that are an expression of that instability. Furthermore, the system (language) is measured on the basis of a dependent variable. The phase transition (e.g. expressed by a temporal instability) of this dependent variable (e.g. the rule of verb inflection) is controlled by one or more independent variables. Time is not a control variable in CT. The use of control variables as a way of finding discontinuities in development is called *modelling*. It is the direct specification of behavioural and control variables in one of the elementary catastrophes.

In the next section, I explain why an analysis based on *control variables* is, in the study of discontinuities of language development, difficult, and maybe even impossible.

Control variables

The application of equations (the so-called catastrophe *analysis*) to (cognitive) development is problematic (and maybe even impossible). This has partly to do with the unknown status of the *control variables*. A control variable is the independent variable of a psychological phenomenon. The behavioural variable is the variable which is measured on a scale or test.

Control variables are needed to observe some of the catastrophe flags. These variables cause the observed change in a system (e.g. language). However, not every independent variable can serve as a control variable. For example, age (or time) increases continually, but age is not a good control variable, because it has been known for a long time that children have individually timed paths of development. Both speed and timing of development vary between children. Van der Maas (1993) has argued that in cognitive development control variables are hard to identify. A possible candidate for a cognitive control variable is memory span which increases over time. In Wimmers, Savelsbergh, Beek and Hopkins (1997), several control variables have been chosen to serve as a guide for motor development. For example, for the development of reaching, muscles, bones, and fat may serve as control variables.

In language development, linguistic variables might also act as a control variables. For example, for the development of multi word sentences, words are needed. So, before a child can lengthen its utterances (or make utterances at all), the child must build up its vocabulary. An increase in vocabulary is a continuous control variable for a potentially discontinuous change in multiword sentences. In the end, an analysis of catastrophes must contain an idea of which control variables guide language development through a discontinuous phase. This leads to an understanding how development is guided by factors that cause transitional change, the same way motor development is controlled by the development of muscles, fat and bones (amongst other factors). See also the appendix for a discussion on equations of control variables.

This thesis on change in language development uses a Catastrophe Theory detection. Since I will apply CT to language development, with the stress on syntactic properties of language, the first question is: how can we apply Thom's theory to (language) development?

Catastrophe detection

Apart from catastrophe modelling (see above) and analysis (i.e. a mathematical analysis of the dynamic equations of a transition processes; this method requires knowledge of the mathematical equations), only *detection* is useful for cognitive development since we do not know what the equations or control variables are that govern a transition (van der Maas, 1993). Detection of a catastrophe happens by using properties or flags that indicate the presence of a transition or catastrophe. The simplest catastrophe model is the cusp and the eight flags.

The cusp and its eight flags

Since the empirically observed data show the level of a variable, not its equilibrium level, a problem occurs with distinguishing a mere increase in growth rate from a sudden jump between two equilibrium modes. Catastrophe theory provides an answer to this problem, by specifying a set of eight empirical criteria that are typical of sudden catastrophic equilibrium jumps. All these indicators, also called catastrophe flags, together mark a transition in a developmental variable.

There are 7 elementary catastrophes (Thom, 1975), which differ in the amount of control variables and the power of the equation (i.e. the more control variables there are, the higher the power is). All catastrophes are sudden changes (i.e. discontinuities) which are caused by changes in control variables. Out of the seven catastrophes, one is frequently applied to cognitive development, namely the cusp. This theory increases the knowledge of processes independent of the field of psychology, and in all, the use of Catastrophe Theory is not limited to the study of (cognitive) development. In social psychology, Tesser and Achee (1994) have studied social interactions using Catastrophe Theory. This is also one of the charming benefits of Catastrophe Theory: it is a formal model that may be applied to any process.

The question arises: why use the cusp, and not one of the other 6 catastrophes? There are three reasons. First, most transitions are controlled by changes in a small number of variables. The cusp demands for 2 control variables. The number of attractor states in the cusp is 2, exactly the number of equilibrium states in the development of functional categories (see next chapter), and it also the catastrophe model with the smallest number of control variables that is needed to explain discontinuities. It is less complex than any other model for discontinuous change. Second, the cusp has proven to be applicable in various areas of psychology (van der Maas, 1993, Wimmers, 1996, Tesser & Achee, 1994). Third, this is only the start of a new area of research using CT. Therefore the simplest model is used.

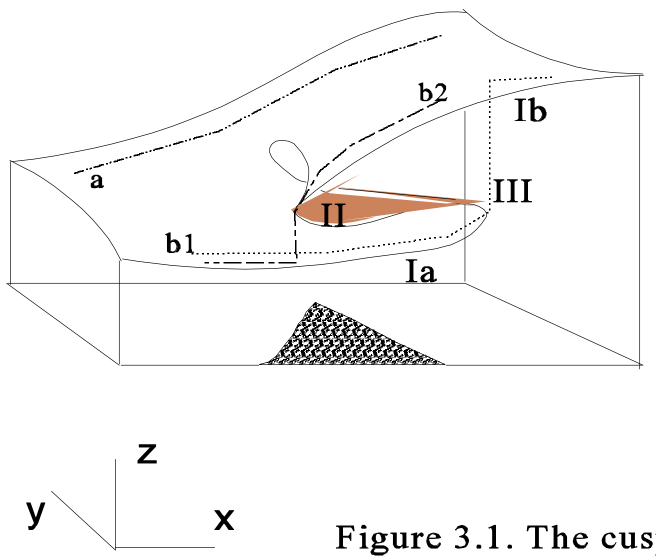


Figure 3.1. The cusp

The cusp looks like a folded table cloth (see figure 3.1) with one behavioural variable (z) and two control variables (x and y). Roughly speaking, we consider equilibria, that is, stable periods in development, and change between these equilibria as the crucial elements of development. The change between these equilibria can take several forms, but in order to decide whether this change is transitional or not, criteria from CT are applied to language development. Following Gilmore (1981) and van der Maas (1993), I relate discontinuity criteria, i.e. the so-called flags, to developmental research. Van der Maas (1993) has found several examples in the literature where bimodal distributions are present. Although such quantitative findings are lacking in language development, it is generally accepted that during the first months, a child changes from telegraphic speech to differential speech (Brown, 1973). These two stable states of speech are a linguistic form of bimodality. It relates to the first flags in CT **bimodality** (**Ia** and **Ib**) which refers to the fact that there are two states (equilibria) in a process. In fact, the bimodality flag indicates the probability distribution of the scores in the transition area. The chance in this area that there are not one but two modes is very high.

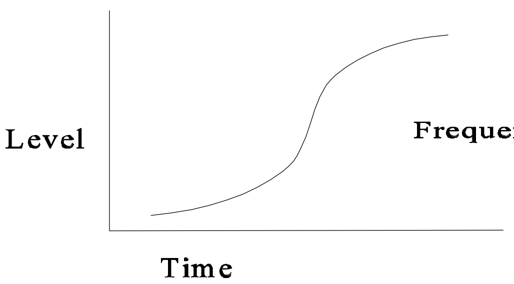


Figure 3.2.1. An s-shaped curve

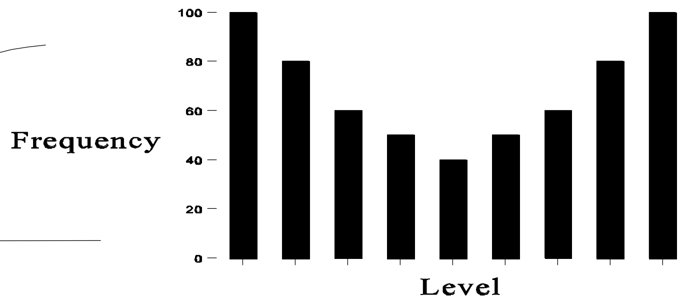


Figure 3.2.2. Frequency histogram of figure 3.2.1.

Van der Maas (1993) has applied this criterion to conservation development. He found a non-conservation stage and a conservation stage. Proposals that there are even more than two states (e.g. a trimodal distribution) are regarded as inappropriate, since, for example, partial conservers are children that are in the middle of a transitional process. Characteristic of partial conservers is a non-constant score on a test during the transition period: one day their score is high (i.e. most answers are correct), the other day many mistakes are made. One must be careful, however, in using this bimodality criterion. Suppose the growth of a test score looks like figure 3.2.1. If the scores are plotted in a frequency histogram, the result would look like figure 3.2.2. Thus, the bimodality flag is not sufficient in order to decide what sort of change is present in development.

Besides that, there is second problem. Namely, how can one test bimodality? Most if not all the widely used statistical procedures assume unimodal distribution. Hartelman (1996) describes a procedure for a statistical test of bimodality. In the transition area, the data are bimodally distributed. Outside the transition area (i.e. the cusp), the data are unimodally distributed.

The second flag helps us a bit further. It is assumed that there is a region that is

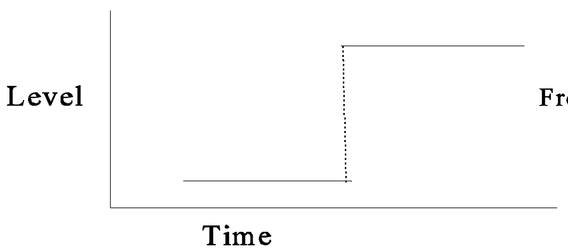


Figure 3.3.1. A sudden jump

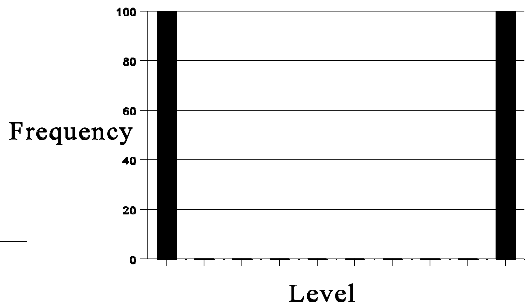


Figure 3.3.2. Frequency scores of figure 3.3.1.

inaccessible (**II**): behavioural variables cannot be in a state somewhere in between these two possible behaviours. This flag prevents the conclusion that a gradual change, in which case there are also two states, is a case of qualitative change. If the data were truly bimodally distributed, the distribution would look like figure 3.3.1 and the frequency distribution would look like figure 3.3.2. However, a problem looms because due to error variance scores may appear in the 'forbidden' inaccessible area. For this reason, inherent to psychological research, it is accepted that there is a probability that scores fall in the inaccessible region (see also Wimmers, 1996). In short, error variance prevents true inaccessibility.

Inherent to the first two flags is a characteristic of change which has been mentioned, although indirectly, by Brown and others. The transition from telegraphic speech to differentiated speech (see chapter 4) has always been referred to as a sudden change. Suppose that change from one state to the other takes place very slowly, would it be possible to decide where one state ends and the next starts? The answer must be no, since there are no sharp boundaries between the states. The only point that can be made is that at the beginning there is a different situation from the one at the end, but this is extremely vague. Apparently the speed of change, in combination with bimodal distributions and inaccessible regions, is important, and although time is not incorporated in CT, the third flag is dedicated to the speed of change. It states that change must occur suddenly. A **sudden jump** (**III**) refers to the fact that a change in equilibrium is abrupt.

A variable jumps suddenly to the second (higher) equilibrium. This is possible at two moments: along the two paths b1 and b2 (see figure 3.1).

The next three flags are indicators of instabilities in the region of a catastrophe. **Divergence of linear response** arises when a system looses its stability after it is perturbed. Perturbations of the control variable(s) near a catastrophe point will lead to a loss of stability, and to large oscillations of the behaviourable variable. Empirically, one needs dense time series (i.e. time series that covers the developmental path of a variable in such a way that every small deviation can be observed) to observe divergence of linear response. The problem in research of cognitive development, however, is that these time series are rare, and the control variable is often unknown. This flag explains why training or corrections in language development seldom leads to success. It has been reported, more than once, that if a parent corrects ungrammatical utterances of a child, the child hardly ever pays attention to these parental corrections. Here follows an example from Braine (1971) (cited in Ingram, 1989):

Child: Want other one spoon, Daddy.
Father, You mean, you want THE OTHER SPOON.
Child: Yes, I want other one spoon, please, Daddy.
Father: Can you say ‘the other spoon’?
Child: Other ... one ... spoon.
Father: Say ... ‘other’.
Child: Other.
Father: Spoon.
Child: Spoon
Father: Other ... spoon.
Child: Other ... spoon. Now give me the other one spoon.

3.1 An empirical example of linguistic stability.

The example from Braine (1971) is not a real example of divergence of linear response, since an empirical example of this flag in language development does not exist. This flag calls for learning or imitation, and that is what is neither expected nor hypothesised in language development. What the example illustrates is that CT predicts that only near a catastrophe point these corrections may have an effect. Likewise, training before and after the transition is very likely to be ineffective. In both cases (before and after a transition), the equilibria are too stable to be perturbed and eventually change to a higher level.

Critical slowing down refers to the fact that a system that is perturbed near a catastrophe point needs more time to get back to its (old) equilibrium. This means that its relaxation time will be longer. This does not mean, however, that the quality of change differs when a system is perturbed: only the time a system needs to get back to its equilibrium is referred to.

This flag is not related to any empirical findings in developmental psychology. However, it is possible to invent an experiment in which a child must repeat sentences in a training session. Suppose a child is in his telegraphic stage (see also chapter 4). One prediction would be that the child will have longer reaction times (eventually leading to non-responses) to the more complex and difficult sentences. Transitional children, i.e. those children that leave telegraphic speech but who have yet to reach differential speech, will show slowed down reaction times, whereas children who have reached differentiated speech, master all sentences without hardly any hesitation.

Anomalous variance is when behavioural variables show an increased variability near a transition. As van der Maas (1993) observed, this has two consequences. First, correlation values will drop near a catastrophe. This means that the structure of the system changes. Old coherences disappear, whilst others have yet to appear. Second, the new equilibrium (i.e. after the jump) may show considerable oscillations that die out when a variable gets further away from the catastrophe. This means that if a child is capable of using a new syntactic rule, the dependent variable might fluctuate during some time around the point of transition.

Like critical slowing down, there is no clear link with existing empirical findings. However, this flag refers to the fluctuations during the transition period which may be found in the score on a test. A second, possible indication that a dependent variable is subject to anomalous variance is when children show inconsistencies in their verbal reports (van der Maas, 1993), or discrepancies between their verbal and non-verbal behaviour. These discrepancies between speech and non-verbal behaviour have been found in children who show in their non-verbal behaviour the opposite behaviour of their language (e.g. acting out experiments; van der Wal, 1996, 108/109). Goldin-Meadow, Alibali and Church (1993) also found a gesture-speech mismatch. They explain this phenomenon as a transitional phase, in which children show a higher error variability in both the cognitive and motor domain. This increased variability points at a loss of structural stability.

The last two flags are vital in deciding the true nature of a catastrophe. The first flag is **divergence**. It occurs when a small change in the initial value of a path through the control plane ultimately leads to large changes in the behavioural variable. In the cusp (figure 3.1), divergence occurs when the control variable changes along the y-axis (from the back to the front) and the behavioural variable has to follow one path or the other. This leads to either the top of the cusp (one stable solution) or to the bottom (the other stable solution). This is what van der Maas (1993) has called the splitting value. If this value is small, there will be a small jump only (or just a acceleration), but if it is high, the jump (as a function of the splitting value) will be large. It must be said, though, that experimental conditions have to be optimal, otherwise divergence cannot be observed. If the setting is not optimal, the result on a test will lead to a unimodal distribution. Again, this shows the strong interdependency of the flags. Divergence shows another feature of the cusp, namely the *bifurcation set*. A bifurcation is literally a “twofork”. This bifurcation set explains why an initial small difference between two paths in the control (i.e. the plane controlled by X and Y) ultimately leads to very large differences (van der Maas, 1993, 14).

Empirically, divergence is closely associated with training effects. The example on the previous page refers to this training effect. As long as the child is not near a point of divergence training will not lead any changes in the language of a child. However, when a child ‘is’ in the control plane, small effects of training may lead to the acquisition of language variables.

The final flag is called **hysteresis**. This flag involves transitions on different points on the time scale, that is, there is a difference in transition points occurring when a control variable either gradually increases or decreases. This can best be explained by following the two paths of b in figure 3.1. The change along path b1 shows a jump. That is, if a variable along the x-axis is increased (from left to right), the variable has to jump somewhere from the bottom to the top. This jump is ultimately on the right side of the tablecloth. The change along path b2 (from right to left) is on the top of the folded tablecloth, but ultimately it has to “jump” on the left side of the cloth. This discrepancy of path b1 and b2 is called hysteresis. Hysteresis has been observed in visual experiments (Ta’eed, Ta’eed and Wright, 1988).

In development, this flag has not been found, although some experiments on conservation show some vague remeniscence of hysteresis (cf. van der Maas, 1993; Boom, Gerlagh & den Hartog, 1997). With respect to language devlopment, categorical perception of sounds like *b* and *p* (in phonological terms: the voiced and voiceless bilabial explosives) is supposed to be relatively abrupt, i.e. with a small fuzzy overlapping area (Eimas, 1985). In a test, when a baby is offered a series of sounds where a *b* changes into a *p* or a *p* changes into a *b* (i.e. in the overlapping area), the perception itself may lead to different jumps, depending on which sound was presented first. Evidence for a difference in perceptual switching can be found in, for example, Bohn and Flege (1993).

In development, this flag can not be observed in longitudinal data, because the direction of x-axis (for example, time) has to be reversed to go from left to right and vice versa. Only in an experiment, it is possible to find hysteresis. Suppose a child is offered 10 sentences that increase in complexity. The child has to repeat the experimentator (a so called EI-test; cf. Ruhland, 1991), and the child is capable of imitating some but not all of the sentences. Somewhere the child will fail to imitate (e.g. sentence 7). If the child were to start with the difficult questions first, his failure to imitate would probably last to another sentence (e.g. sentence 3). Although longitudinal data are used in this study and therefore this flag cannot be found, hysteresis could be found with such a test.

An important concept of CT is structural stability. Perturbations of a system under developing constraints will not lead to major reorganisations. The system (i.e. language) develops without any reorganisations or instabilities, except for that point in time where the system undergoes the transition. Structural stability means that empirical findings are similar in repeated conditions and small changes in these conditions do not alter the results dramatically. Another aspect is that time is not explicitly formulated within CT. CT is a local theory and model. Changes may be catastrophic, and still take place over a longer period. For example, Zeeman's dog changes its mood within seconds (Zeeman, 1976), whereas change in children learning the process of conservation takes place over 2 years (van der Maas, 1993). Nevertheless, the jump itself (on the underlying equilibrium level) is instantaneous.

Evidence for catastrophes in (developmental) psychology

I have dealt with one of the seven catastrophes, namely the cusp. The eight flags that mark the cusp can be linked to the behavioural phenomena in general and to psychological development in particular. Zeeman (1976) modelled the instantaneous behaviour of dogs when they were brought to distress. There is no continuum: dogs suddenly change from attack to retreat or the other way around. The two equilibria, attack and retreat, have two control variables: fear and rage. There is no intermediary state of behaviour. The dog either attacks, or retreats. The two modes, the lack of an intermediate behaviour, and the suddenness of behavioural change make Catastrophe Theory very suitable for explaining the empirical findings. Other studies have found catastrophic change in psychological processes. Visual perception, i.e. the perception of multistable figures, has been modelled with Catastrophe Theory (Stewart & Peregoy, 1983).

Van der Maas and Molenaar (1992) report on a study of conservation development. Several indications have puzzled researchers over the years. Children do not have the ability to apply conservation before the age of (roughly) 5. Within 2 years, a child learns how to conserve quantity, length, etc. Van der Maas and Molenaar show that this process is characterised by catastrophic change, since there is a sudden change (sudden jump flag) from a non-conserving to a conserving state (bimodality). Furthermore, they found inaccessibility and anomalous variance. Van der Maas and Molenaar have reformulated Piaget's equilibrium theory into a formal mathematical theory. Piagetian problems like the lack of a stage and a transition definition have been solved.

In the field of motor development, Wimmers (1996) has performed a study in the development of reaching and grasping, showing that the transition from non-grasping to grasping is a catastrophe-like change. In sum, these studies reveal that the application of CT helps to model and to understand the nature of (non-)developmental processes in psychology.

Catastrophe theory provides clear criteria to determine the presence of discontinuities. There are eight flags, which can be linked to empirical research. CT is a formal (i.e. mathematical) model of phase transitions, that like any other formal model of language development, needs a structural model (see chapters 2 and 4).

3.3 Summary and conclusions of the chapter

One of the fundamental problems in the study of human development that still awaits a solution is the proof of psychological discontinuity (see also chapter 2). In the past, there have been no clear definitions of discontinuity and no clear criteria to demonstrate discontinuity, either. A special branch of mathematics called non-linear models is helpful for fitting and modeling development. These models can describe behaviour processes that show chaotic, non-linear continuous or discontinuous behaviour. That is, behaviour that does not follow smooth paths along the time continuum.

Dynamic systems models are helpful in four ways. First, empirical data can be *fitted* to decide what sort of change one is dealing with (i.e. a help in deciding wether or not change is non-linear). Second, they incorporate and explain the *mechanisms* responsible for change. Third, a *model* can be made to describe development. Fourth, *predictions* can be made on the basis of the model.

There are two types of models and theories in this branch of mathematics, both capable of explaining several characteristics of development. First, there is a group of models called *dynamic systems models*. These models are based on non-linear (differential and difference) equations. The models are reductions of *individual paths of development* (thus, not averages of groups), and it is assumed that cognitive growth is constrained by limited resources (e.g. memory, attention and motivation), that there is an end state of the cognitive system which is called the carrying capacity, that cognitive growth is hampered by what is called feedback delay (i.e the effect of a specific growth level on a later growth state), and that there are supportive, competitive and/or neutral relationships between growers. Cognitive growth is expressed in a number of parameters in the equations. There is an initial state, L_t. K is the carrying capacity, that explains the final state. A variable (or grower) grows from L_t to K. The process depends on a growth rate r that controls the speed of change. The exponent parameter p controls the power of the equation. This parameter explains to what extent the current level of a grower specifies the further growth of L.

Catastrophe theory is a formal theory of discontinuous behaviour. One of the seven elementary catastrophes, the cusp, proved to be very useful for the detection of catastrophes. Eight flags together indicate a catastrophe. Some of them (a sudden jump and bimodality) are not new in research on development. Catastrophes are characterized by instabilities between two modes (or equilibria), a rapid change between these equilibria, and perturbation of the developmental system during the catastrophe leads to either a new equilibrium or to a labourious restoration of the old equilibrium.

The research questions may now be reinterpreted. In chapter 2, I argued that theories on language development lack a thorough formal model for the proof of assumed discontinuities processes in development (i.e. discontinuity in the psychological sense; cf. Piaget's equilibrium theory). The solution for the lack of models for the study of discontinuous behaviour is a branch of mathematical models that are called non-linear models, i.e. continuous non-linear and discontinuous models. They test the non-linear assumption(s) of development. The study of discontinuities in development has profitted by Catastrophe Theory, which provides us with criteria that can demonstrate discontinuities in development. On the basis of Catastrophe Theory, the central research question now reads:

> Is there evidence for continuous or discontinuous (both in the psychological, i.e. quantitative, meaning) development as defined within the linear and growth model paradigm (continuity) or as defined by the catastrophe flags to language development (discontinuity)?

The next chapter contains the empirical basis of this study, which consists of the children that participated in the study, the linguistic variables, the method of analyses, and the length and frequency of recording and the transcripts.

4

The bare necessities: the method of the study

And go along not thinkin' about it
I'll tell you something true
The bare necessities of life will come to you

The song **The bare necessities** from the Disney movie **The jungle book**

4.0 Introduction

Problems with respect to discontinuous development have been discussed in chapter 2. The problem is that traditional approaches do not have clear criteria to decide for or against non-linearities or discontinuities. I therefore introduced dynamic systems theory and catastrophe theory. The latter theory can explain discontinuous change (chapter 3).

In this chapter I zoom in on the variable of this study. The goal is to find empirical evidence for discontinuities or continuous nonlinear change and to answer the main research question in chapter 5. The choice of the variable(s) should satisfy two criteria. First, the empirical variables have to make up an important aspect of language. Second, the variable(s) should have the possibility of sudden change. That is, a variable should develop in a potentially rapid or even discontinuous fashion.

In the vast body of literature on language development most papers and books deal with structural (e.g. syntactic) analyses. In my experience, quantitative change plays a minor part in the study of language development, but some of these quantitative findings of Brown and others (e.g. Wells, 1986) is reviewed, with an emphasis on the presence of possible discontinuities and stage wise development. The variable of the study, *function words*, is introduced in section 4.1. I discuss the general characteristics and the linguistic properties of function words (i.e. the structure of function words) on the basis of three examples (i.e. pronouns, modals and articles).

In section 4.2 I introduce the conditions for the study of catastrophes in language development, i.e. the quality of the data, the use of frequency (e.g. of the recordings) and the length of the recordings.

Six children have been recorded for the analysis of their speech. Their age, the length of the period of the recordings and the transcripts of the children's utterances, and the programs from the CHILDES project are presented in section 4.3.

The summary (section 4.4) describes the tools that are needed for analyses of language data of children on the basis of catastrophe theory criteria, and a reformulation of the research question.

4.1 Function words in language and development

Introduction

The first part of this chapter deals with the structural properties of function words. The relationship between function words and linguistic theories is stressed (cf. Ruhland, Wijnen & van Geert, 1995). The second part of this section discusses the development of functional categories. Linguistic theories and their prediction on the development of function words are discussed, and I introduce three specific cases of function words, namely articles, pronouns and modals. The syntactic properties of these words are discussed, and I emphasise research on the *development* of these three function words.

Function words in general

For syntactic reasons linguists make a distinction between lexical items and functional phrases. Lexical phrases consist of lexical items such as nouns, verbs, and adverbs. In the early stages of development, prepositions are also assumed to be lexical, whereas in later stages of development, this class of words is used in a syntactical way. Lexical items are the first acquired words, and in most cases they refer to the surrounding physical world. That is, content words refer to a non-linguistic reality. The concepts can be regarded as references to what we see, hear, do, smell, etc., in short, to all entities and actions in our world. Functional phrases are more or less abstract words that have a grammatical rather than a referential use. These phrases, directly related to function words, are built up from, for example, articles (*the* and *a*), modal verbs (e.g. *must*, *may* and *will*) and pronouns (e.g. personal pronouns like *I*, *you* and *we*). Function words are called closed class words because they form a class that is limited and that is hardly ever expands, unlike lexical items which are expanded each time a new word is coined for a new cultural phenomenon. For example, the noun *computer* has been added to the vocabulary of many languages in the world since the 1970's, but articles are a class of words that have not been expanded for a long time: in Dutch there are two definite articles (*de* and *het*, ‘the’) and one indefinite article (*een*, ‘a’). Furthermore, function words have some intriguing characteristics. Apart from the aspect of constituting a closed class and the relationship with syntactic structures, function words have systematic distributional properties in sentences, i.e. they appear in fixed positions, they often are monosyllabic and compose a small set of phonemes, they receive weak stress and may undergo cliticization, i.e. be attached to other words, and they are extremely frequent in speech (see also Gerken, Landau & Remez, 1990, Ruhland 1997). Note that *function words* are not the same as *functional categories*: functional categories refer to a syntactic analysis, whereas function words refer to the surface form of these categories.

Given their primarily grammatical function, the productive use of function words requires a relatively high level of syntactic differentiation, and vice versa. The frequency with which function words occur in the child’s language constitutes an important indicator of syntactic development. Function words also appear relatively late (despite their frequency and systemacy in the input and their monosyllabic character) in development, i.e. when sentences have a MLU (i.e. Mean Length of Utterance) of over 1.75 (Brown, 1973). The appearance of function words constitutes a new stage in language development, namely the differential stage which follows the telegraphic stage. The structural difference between these stages lies in the use of function words, inflection etc. during the differential stage.

With respect to syntactic trees (chapter 2), functional categories are associated with syntactic features such as tense, agreement and case. These categories form a ‘shell’ over the lexical categories. This shell is an interface between the representation that specifies the thematic relations between the verb and its argument noun phrases, and phonological form. Function words form one subclass (apart from, for example, verbs inflection) within this class of functional categories.

Function words are typical in a linguistic way. They are grammatical markers in a sentence, and they have a function in relation with other words (i.e. lexical items). They refer to other words, phrases or sentences. For example, *he* in example 4.1 refers to **the man**, and *her* refers to **his wife** (this is indicated by the use of the index i and j).

4.1 The man$_i$ is unfaithful to his wife$_j$, because he$_i$ is married to her$_j$.

These words are syntactic elements in a sentence, they do not refer to definite entities in reality. In other words, they are related to functional categories in a language, i.e. associated with syntactic structures. Their distributional properties in sentences are systematic: they appear in fixed positions. Articles appear in front of nouns, not behind them. Furthermore, they often are monosyllabic and constitute a small set of phonemes.

They receive weak stress and may undergo cliticization (i.e. be attached to other words). Examples in Dutch are ‘kep’ (*ik heb*, I have) and ‘daggie’ (*dacht hij*, thought he). They are extremely frequent in speech, and additionally, there is no dependence on the subject of a conversation. Function words limit the meaning of contents words. *My guitar* refers to a much more limited class than *a guitar* or *guitar(s)*. Function words in Dutch have been summed up in van Wijk and Kempen (1980). There are 418 function words in Dutch, classified in 8 categories: prepositions, conjunctives, determiners (adjuncts before a noun), relatives (linking relative clauses with their antecedents), pronouns, auxiliary verbs, adverbs, and adverbs of degree (qualifiers). There are only few data on function words (in development or in the speech of adults) with respect to frequency.

In sum, the reasons for choosing function words, and not, for instance, all words that end with an e or an x, are that this class of words is syntactically important, that they are not dependent on the subject of the conversation (and therefore, quantitative aspects important in the analyses are not hampered by the problem that there might be too few occurrences in the language of a child) and that they constitute a closed class group of words. The combination of these three aspects makes them a very interesting and linguistically and communicatively important group of words.

The question is what one may expect with respect to developmental sequences of functional categories as compared to lexical categories. There are three general hypotheses about the development of syntax. All three make a claim about the origin and the structure of the initial competence of early child language.

Hypotheses of development: full, reduced and no competence

Three approaches make a claim about the possible nature of underlying syntactic structures in development. These approaches to structures in language (development) are based on linguistic theories (cf. Ferdinand, 1996, Ruhland, 1997). First, there is the hypothesis of *full competence* (FC). According to FC, children have full knowledge of grammar structures with birth. FC is based on the continuity assumption: all UG options are available to the child from birth on. Change is only a behavioural (i.e. superficial) change, underlying structures are fixed. However, children leave out functional elements in their actual language production. Solutions to this problem are reduction rules, production limitations, and phonological development (Ferdinand, 1996).

Second, the *reduced competence* (RC) hypothesis leaves a few blanks in the initial syntactic knowledge of a child. Clahsen, Penke and Parodi (1994), for example, assume that the functional structure of a sentence is extended gradually.

Third, the *lexical thematic* (LT) hypothesis states that all relationships between words and phrases are non-syntactical at early acquisition stages. LT is problematic on a theoretical level since it is unclear how “lexical thematic children” reach syntactic knowledge since this knowledge is not provided by the parents. FC is problematic on a behavioural level when children do not show any overt command of syntactic knowledge, which they should have according to FC.

Since there are three assumptions of the origin of functional categories there is not one position with respect to the change that may come about, and independent from the point taken in the discussion on the nature of a child’s knowledge of syntactic structures, a theory on development needs to express, apart from the acquisition of structural properties of function words, the expected change in the development of function words as functional categories.

Theories on the development of functional categories

In the initial phase of language development, children first use proper names (most notably, their own name, instead of ‘I’), or common nouns that can be interpreted as such, instead of pronouns. After a period in which determiners, such as the definite articles *de* and *het* (‘the’), and the indefinite *een* (‘a’), are not realized at all, children start using *schwa*s in the appropriate positions (schwa’s are fillers, an ‘uh’-sound). Modals are used, but only in the present singular. The common denominator in these aspects of syntax and syntactic development is the concept of ‘functional projections’ (see chapter 2). Determiners (e.g. articles) and determinative pronouns, for example, are considered to be the head of the DP (Determinator Phrase). If pronouns and determiners are systematically absent, this may mean that the grammar lacks the category DP. Note that the absence (or unsystematic use) of particular elements that are associated with functional categories does not necessarily lead to the conclusion that these categories are absent (see Hoekstra & Jordens, 1994, for a critique of these assumptions).

Three theories on language development predict distinct quantitative increases in language development. Radford (1988b, 1990b) has repeatedly argued that during the initial phase of combinatorial speech, grammars lack functional categories and that the associated modules of grammar (the Case module) are inoperative. At this time, only the ‘categorial’ modules (e.g. the lexicon) are active. The X-bar module (i.e. the set of principles governing the construction of hierarchically structured phrases from lexical entries) is also available, but the set of categories it can operate upon is limited to the four lexical classes Noun, Verb, Adjective and Preposition. The transition from this limited grammar (subgrammar) to the target grammar has yet to be determined by the child. This target grammar comprises functional categories and the associated principles of grammar. *Maturation* could induce the transition. It is conceivable that the modules of grammar that are associated with functional categories need more time to mature than the principles that govern lexical categories and their projections. All functional categories are interrelated, they address the same modules of syntax (notably the Case module)). Therefore, they are expected to emerge at the same time.

An alternative hypothesis is that all functional categories and the associated grammatical modules are available from the start, but that a certain amount of input (i.e. specific lexical and morphological items) is needed for them to be 'triggered' (i.e. become operative). Clahsen, Eisenbeiss and Penke (1994), for instance, suggest that acquisition of the regular subject-verb agreement paradigm results in the incorporation of the corresponding grammatical features into the child's phrase-structure representation. "This may lead to new phrase structure layers or to the specification of previously existing functional positions" (Clahsen, Eisenbeiss & Penke, 1994: 18). This is argued to be evidenced by, among other things, unrestricted movement of finite verbs. They furthermore contend that the emergence of DP is a consequence of the acquisition of inflection, since both relate to the same syntactic (case) features (see also Chapter 2).

The main difference between the maturational account and the *Lexical Learning* (LL) approach of Clahsen and his collaborators is that the latter does not imply a qualitative shift in syntactic development. In LL, there is no change in the carrying capacity (see chapter 3), the body of syntactic knowledge, associated with either the growth of finite verbs or pronouns and determiners. Clahsen, Eisenbeiss and Penke's (1994) suggestion entails a process of learning by generalization. Their proposal seems to imply that the acquisition of the agreement paradigm is a necessary condition for the acquisition of pronouns and determiners, which may be called a precursor relation (between finite verbs and pronouns and determiners). The quantitative prediction is a gradual increase, and maybe even a linear change. Sudden changes are not expected from this theoretical assumption.

With respect to the relation between function words and *Parameter Setting*, language development is a question of pushing the right buttons. Parameter setting, in its simplest form, is a theory that assumes that all linguistic knowledge is present from the beginning (a continuity approach to language development), but it takes time to switch all buttons to their correct state. This delay in time is the result of other linguistic factors, as well as psychological factors. The child's task is to determine which of the core grammar variants allowed by Universal Grammar generates the sentences it is exposed to, and to set the values of the parameters accordingly (Chomsky, 1986, Gibson & Wexler, 1994). Parameter setting was conceived as a solution to the logical problem of language acquisition: it explains the possibility of acquiring a complex grammatical system in the absence of sufficient data. It pertains to an idealized model of acquisition, which assumes simultaneous presentation of all relevant linguistic stimuli (triggers) and an instantaneous, non-graded transition from the initial state - UG in its pure form - to the final state - the target grammar. It does not need much contemplation to see that language acquisition in reality can never be instantaneous, even if it is assumed that the child is a perfect parameter setter, simply because the full array of necessary triggers can not be presented simultaneously. Still, acquisition of grammar might take place within a very short time. Given a sufficiently large passive vocabulary, a relatively small number of input utterances would be sufficient to set all parameters. However, on the assumption that the order of presentation of triggers is non-deterministic, the order in which parameters are set, or, in other words, the sequence of intermediate grammars would be unpredictable. Finally, it is conceivable that all principles are available, and that all of the necessary triggers are accessible, but that particular parameters cannot be set until certain other parameters have been set first (*ordered parameters*, Weissenborn, Goodluck & Roeper, 1992; see Verrips, 1991, for a critique of this perspective).

However, as pointed out in Ruhland, Wijnen and van Geert (1995), sudden changes of an underlying covert variable do not necessarily have to lead to abrupt changes in the overt variable. A structural change in some underlying mental capacity may fail to produce an abrupt change (discontinuity) in behavior because the behavioral reflexes of the new state may be in competition with the old. The net result of this competition would be a gradual adaptation of the performance system to changes in competence. Anyone coming from the Continent must have felt the sensation of walking through a Scottish city (or in any other left driving country) and planning to cross a road. In most countries one has to drive on the right lane, and cars come from the left first. In a Scottish town the cars come from the right first. The absence of control over this situation is dangerous, but the dangerous situation will die out if one stays long enough in one of the left-driving countries. This “old habit lingers on”-principle may explain why a structural change may not instantaneously lead to a new equilibrium (i.e. an direct adaptation to the new situation). Possible quantitative predictions include a sudden change and a gradual but rapid change to the new equilibrium. However, a precise prediction is not possible.

With respect to functional categories and function words, predictions of the quantitative development of function words on the basis of linguistic theories are diverse. Learning theories (like Lexical Learning) would predict a non-stagewise change. Change can impossibly be discontinuous. Parameter setting may explain and predict sudden changes, as all “switch” theories predict. Maturational accounts of development (in the true sense of the word maturation) are growth theories, which predict gradual change.

The conclusion must be that theories on syntactic development contain a limited explanation for the way syntax develops in sense of psychological (dis)continuity. Existing structural (linguistic) theories allow for both continuous and discontinuous predictions (in psychological meaning). Nevertheless, these theories do not predict language development in terms of exact change. Here it is assumed, in accordance with a number of authors, that functional categories develop, i.e. not all functional categories are present from the beginning (De Viliers, 1992, Ferdinand, 1996). However, this structural assumption does not predict the quantitative paths of development. The way these paths develop in a quantitative way has been studied by Roger Brown (amongst others). His work and related work is reviewed in the following section.

Function words development: general and specific observations

In a series of articles Brown (1973, 1976; Cazden & Brown, 1975) has shown that the development of function words and functional morphemes has a characteristic pattern. They are suddenly introduced in the language of a child. Take a look at the following graph.

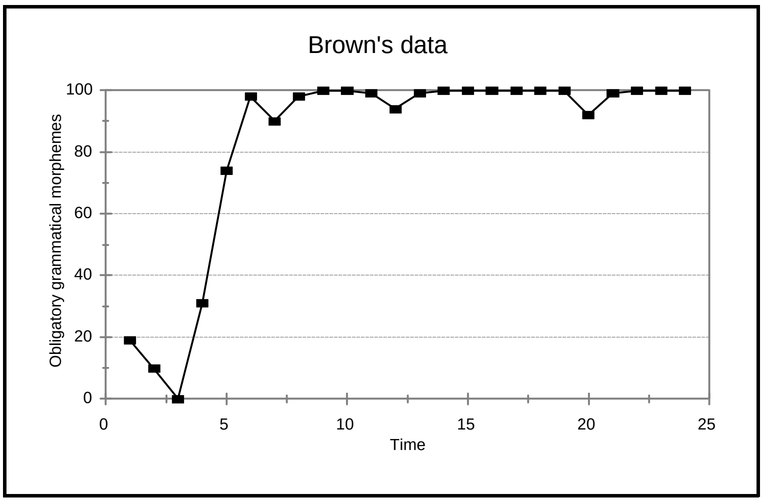


Figure 4.1 The development of the plural morpheme (Brown, 1973);
Age range from approximately 116 to 208 weeks.

Within months, the child (here Sarah) in Brown (1973, 256) shows a rapid development in the acquisition of the use of pluralization. For this reason two stages have been distinguished. Before the introduction of function words the stage is called telegraphic speech. In this stage sentences have limited length and there are no function words. After the introduction the stage is called the differentiation stage. Sentences become longer and they contain function words. There is a normal word order in both stages (i.e. no violations of the grammar prescriptions of a language). Quantifications or growth curves of grammatical morphemes like personal pronouns and articles lack in Brown's work (apart from two graphs in Brown, 1973), but Brown's main contribution to language development is that he developed a coherent framework of child language from the early stages to the acquisition of the grammatical morphemes, showing that the development of these morphemes is rapid. Cazden and Brown (1975) explicitly say that their findings only apply to English. According to Schlesinger (1971), however, telegraphic speech (or telegraphese) is observed in all children in all languages. Cazden and Brown (1975) claim that determinants of development are not frequency of the input nor the salience of function words. The development of function words is determined by their cumulative complexity (i.e. the various complexity of function words explains their development). Finally, the acquisition of syntax is impervious to deliberate assistance. In terms of catastrophe theory: the system has a structural stability.

In the mid eighties, a comprehensive, quantitative study of function words was published by Wells (1985). Over 100 children participated in the study. They have been measured once a month. The variables Wells studied overlap with those in this study, namely function words.

In figure 4.2, I present the data of Wells study. One can see the beautiful s-shaped form which is rather characteristic of large groups of subjects, measured and averaged over age. The analysis of Wells' graphs shows that the shape of figure 4.1 is indeed an s-shaped curve (Ruhland, 1996).

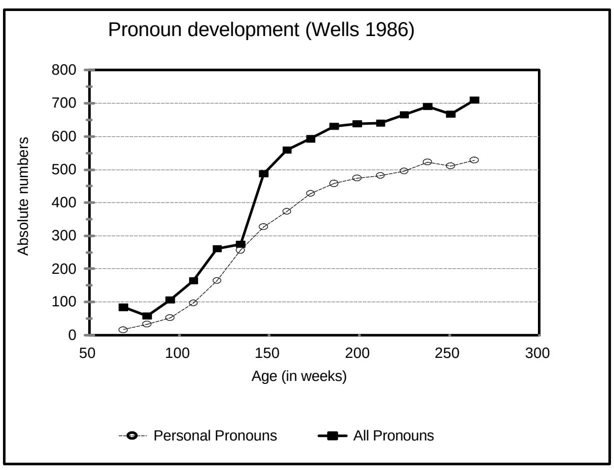


Figure 4.2. The development of pronouns.

What is clear from the quantitative study of Wells is that monthly cross-sectional recordings yield near perfect s-shaped graphs. However, this is not in line with Brown's findings, i.e. the sudden change of function words. These group studies with large recording intervals are not suitable for the study of development since they average out individual trajectories, and they leave out the intermediary changes within the development of function words.

Pronominal vs. Nominal style

Qualitative differences between children in early language development (i.e. the development of function words like pronouns) are based on the notion that children have different styles. These styles have been called *nominal* and *pronominal* (Bates, Bretherton & Snyder, 1988, Bloom, Lightbown & Hood, 1975, Shore, 1995), or *expressive* and *referential* (Nelson 1975). Bloom et al. (1975) found that if MLU equals or exceeds 2.5 (i.e. an average of 2.5 morphemes per utterance), the variation in children decreases. With shorter sentences (i.e. an MLU smaller than 2.5), they conclude, a child's first sentences are either pronominal (using mainly pronouns) or nominal (using nouns), and these two styles are not mutually substitutable in the beginning. Children use either a pronominal or a categorizational strategy. Pronominal and nominal variation explains the telegraphic and pivot contradiction since different children can travel different paths, but they reach the same end. In addition, Nelson (1975) found that in the language of referential children (R), nouns decline as MLU increases, whereas in expressive children (E) nouns increases as MLU increases. R-children start with a high proportion nouns to pronouns, and their developmental course amounts to an increasing use of pronouns. E-children start with a balance of noun-pronouns, and they become increasingly lexical. E- and R-children developmental patterns are not discrete types but different points along a continuum of greater or lesser lexical-syntactic emphasis.

The question is whether the distinction between pronominal and nominal (Bloom, Lightbown & Hood, 1975) or referential and expressive styles (Nelson 1975) enables us to make predictions on the quantitative change we might expect in language development. We think that there is no compelling reason to believe that a linear, a slow gradual or a very rapid change follows logically from either of the styles that according to Nelson (1975) and Bloom et al. (1975) exist in language development. Pronominal children may acquire pronouns early in their language development, but this development may be either very rapid, slow, or linear. In a similar reasoning, nominal children are late, but when they catch up they may show any quantitative behaviour (from linear to sudden change). The style of development, so to speak, does not rule out any of the kinds of change proposed. My analysis will shed a light on the issue of two (or more) paths in terms of quantitative patterns.

Provisional conclusions and summary

Function words relate to several linguistic and developmental theories in the sense that these words call for a high level of syntactic knowledge. I discussed three hypotheses on the linguistic patterns of change. First, it is conceivable that the modules of grammar that are associated with functional categories need more time to *mature* than the principles that govern lexical categories and their projections. An alternative hypothesis is that all functional categories and the associated grammatical modules are available from the start, but that a certain amount of input information, specific lexical and morphological items, is needed for them to become operative. At the surface level, this suggestion entails a process of *learning* by generalization. *Parameter setting* assumes that all linguistic knowledge is present from the beginning (a linguistic continuity approach to language development), but it takes time to switch all buttons to their correct state. Predictions on the quantitative development of function words on the basis of linguistic theories are very general. Maturational accounts of development (in the true sense of the word maturation) are growth theories, which predict gradual change. Parameter setting may explain and predict sudden changes, as all “switch” theories predict. Learning theories (like *Lexical Learning*) would predict a non-stage wise change.

It is assumed that since all functional categories are interrelated (i.e. they address the same modules of syntax) they are expected to emerge at the same time. I might paraphrase The Beatles’ song *Come together* here: *All together, right there*. “All” are function words and “there” is the third year of a child’s life. The development (i.e. the introduction) of function words is a process in which several function words develop in the same time span. The interrelationship between function words is supposed to be very high. This interrelationship consists of the syntactical and distributional properties. Finally, the quantitative behaviour of syntactic elements like function words is predicted as a very general one, namely that they are suddenly introduced in the language of a child. This study aims, among other goals, at filling the gap of a method that uses the data of frequently recorded children.

From this summary, four developmental hypotheses about function words are drawn. The four hypotheses are in fact specifications of the main research question in chapter 1.

1. Function words address identical syntactic slots, i.e. functional categories.
2. Therefore, function words are qualitatively highly related language variables.
3. As a result, function words are assumed to be related in quantitative shape and in timing (i.e. the timing in the developing language, not as an age related phenomenon) between the children.
4. Function words appear suddenly as the result of their closed class characteristics (they come together because they are not separate, but related words).

This study aims, among others goals, at filling the gap of conclusive empirical evidence. Suppose that function words appear suddenly in child language. What is needed to perform quantitative analyses, in order to find evidence for a sudden (i.e. catastrophic) change?

4.2 Methods of analysis: frequency analyses

Introduction

Traditionally, linguistic approaches to development emphasize syntactic acquisition. Pronouns, for example, have not been counted, but these words were subjected to structural analyses (using a small number of examples) (cf. Koster, 1993). Quantitative analyses of linguistic structures are not commonly used as a method of empirical research. However, for the analyses of data and in search of catastrophes, a quantification of data is needed. This quantification calls for a method that reliably allows for conclusions on the nature of change in development. This method (i.e. conditions) differs from traditional research, especially with respect to the density of the time series (see also Ruhland, 1996).

Conditions for frequency analyses

Frequency analyses of time series in language development research must satisfy a number of conditions. In a number of recent papers (Behrens, 1993, Wijnen, 1996, Ruhland & van Geert, 1998a), this kind of analysis has been used to study language development. Part of the research with frequency analyses is exploratory since there is no fixed idea about, for example, how to determine the density of measurement. There is a need for more specification, and it is essential to formulate a number of guidelines or conditions, based on general principles and empirical analyses. So, the question is: what is needed to perform frequency analyses in language development? The obvious answer is, of course, language development data. However, it is not as simple as that. In the literature some general principles can be found with respect to the repeated recording of child language (I discuss longitudinal research of spontaneous language data only). Ingram (1989), for example, states that 'the child is visited at predetermined intervals for a reasonable length of time with the purpose of collecting a representative sample.' Three concepts are important, namely *intervals, reasonable length of time*, and *representative* (I will use the word *standardized*, since representative refers to a sample from a population).

The length of the interval between two recordings must be short enough to cover the developmental path (i.e. the shape of change). In many cases, an interval of two weeks has been considered to cover language development sufficiently. Some researchers employ a 'one day sampling' schedule (Brandenburg, 1915). Every word, every sentence, spoken by a child should be recorded. The ideal situation with respect to interval and the length of recording would be something like the Seven Eleven stores, which are open 24 hours a day, 365 days a year. One would be able to see the variance of a child during a day or any other time index. However, one can see that this is nearly a mission impossible. The amount of time spend on collecting, transcribing, and analysing would be nearly infinite. Furthermore, the question is whether day to day sampling yields more information than recording every week or every two weeks. The density of the intervals of recording depends on another factor: what language variable is one interested in? Some language characteristics appear rather slowly in the language of a child, while others come rushing in. Thus, the frequency of recording also depends on the sort of variable. With respect to syntactic variables (as in this study), two-weekly observations was my starting point in order to cover the developmental path. Follow up studies must be carried out to yield a better understanding of the recording density needed.

A second aspect of recording child language is the length of the recording session. This length of recording has usually been one hour (cf. the Childes Database for examples; see also Nelson, 1973, Wijnen, 1996). It might, however, very well be that a representative sample is obtained only after two or even three hours of recording language development (see also the study of emotion; de Weerth & van Geert, in press). The question is again: representative of what? The recordings should cover so many utterances that an increase of recording time no longer changes the (average) score of a certain variable. Thus, if the average amount of, for example, two word sentences in one hour of recording is the same as the average amount in two hours of recording, we have a representative sample using one hour of recording. Of course, taping half an hour could be sufficient as well. So, no clear-cut proposals have been made with respect to this 'length of recording'. In this study I used one hour as a probably sensible index of syntactic development (see also Nelson, 1973, Wijnen, 1996).

Lastly, the quest for a standardized sample means that if children are to be compared with each other, the situation in which the children have been recorded must be exactly the same. The same toys, the same house and even the same mother. This is, of course, impossible. What is possible is a general setting: the mother is present, the child is playing with toys, no guided questions, and the child is left in her daily routine. In my study, this general setting is kept as much the same as possible. Every recording of a child has been made on the same moment of the day, the child is playing with toys, and the mother is present. Differences between sessions (in terms of, for instance, the amount of words spoken) are individual variation and/or error measurement.

From a practical point of view, I use an interval time of two weeks for recording longitudinal spontaneous data. In recent papers, this time series and frequency scores have been used to study the development of verbs in German and Dutch (Behrens, 1993 and Wijnen, 1995), and to study the development of function words (Ruhland & van Geert, 1998a). Observations with a **two week interval** of **one hour recording** in a **constant setting** are used in this study.

The three concepts, i.e. the time interval, the length of recording and a standardized setting, are important because frequency scores are a measure of the *use* of a variable in the language of a child. This use is the overt form of a covert syntactic rule. Frequency scores of a variable measured over time only are not a process. However, through theories like dynamic systems theory or catastrophe theory we can interpret these time series of frequency scores as the overt form of a process. Additionally, a linguistic theory is needed to explain the process in terms of relationships between sessions (the process) and between variables (structure). These qualitative analyses are used to show that frequency scores are more than just a number: the amount of, for example, articles are indicative for the amount of DP’s (Determinator phrases), and therefore, frequency scores refer to both the acquisition of grammar, and the process of acquiring that grammar.

Proportional or absolute, that’s a question

One objection to the use of frequency scores could be that frequency scores based on absolute scores do not highlight a process, but that they are a collection of random numbers (i.e. measurement errors). That is, it is argued that absolute scores reflect only partially the true nature of development, but that variance and measurement errors obscure the ‘real’ scores, and therefore these absolute scores are rather random numbers. Due to measurement errors and variance these frequency scores of absolute numbers would not be representative for the development of, for instance, language structures. Of course, the question is if proportional numbers should be used. An example will show that this is equally problematic. Suppose that a child is not using any variable (e.g. the number of inflected verbs) until a specific age. The frequency score of that variable is 0. Then, suddenly, the absolute level of inflected verbs jumps from 1 to 20 cases (a graph of this increase would look like figure 3.3.1 in the previous chapter). Suppose furthermore that an independent variable is changing suddenly, for example, the number of words (i.e. all words spoken during a session) increases from, let us say, 10 to 1000 in the same age period. As a result, the proportional number of inflected verbs increases from 0.1 (verbs divided by words = 1/10) to 0.02 (verbs divided by words = 20/1000). Proportional and absolute numbers differ by their true nature. In the example given, there is a decrease in the proportion, but an increase in the production.

Behrens (1993) used corpora of child language to study the temporal reference in German child language (i.e. the development of verbs). One of her findings was that sentences with a verb, as a proportion of the total amount of sentences, increase quite suddenly. This is a remarkable finding for two reasons. First, not only function words (according to Brown and others) develop suddenly, but so do sentences with a verb (i.e. all sorts of verbs, finite and non-finite). Second, proportional scores do not rule out the possibility of sudden change. Wijnen (1995) showed on the basis of time series of recordings of the same children as in this study that the development of finiteness in Dutch is a gradual process. Children acquire finite verbs first slowly, then the speed of acquisition increases, and finally, the process slows down until all verbs are finite (in all cases where finiteness is obligatory). Both Behrens and Wijnen show that proportional frequency scores are optimally suited for bringing about the form of change. This form of change has not so much to do with grammar, but it concerns the mechanisms that have to be associated with the development of verbs. The most important finding, however, is that these proportional frequency scores show changes, sudden changes being one of them. Proportional scores show the importance of the dependent variable as a subset of the overall structure or system (e.g. words, sentences, or language in general). The use of proportional numbers, however, is not free from problems. The disadvantage of proportional numbers is that small absolute differences can lead to big proportional differences. To circumvent this problem, absolute numbers are used in the analyses. In chapter 5, these absolute numbers were checked for possible influences of other, independent variables.

Provisional conclusions

Most important criterion to use frequency scores was the need for quantitative data. These data are used for analyses on the basis of catastrophe theory and dynamic systems theory. The conditions or guidelines for frequency analyses of time series in language development must satisfy three elements:

a. intervals that cover the developmental path;
b. a reasonable length of recording time;
c. a standardized session.

Two-weekly observations of one hour in a general setting (the mother is present, the child is playing with toys, no guided questions, and the child is left in his daily routine) were used as the starting point in order to cover the developmental path. Frequency scores are a measure of the *use* of a variable in the language of a child, i.e. the overt form of a covert syntactic rule.

With respect to the objection of using absolute scores, I argued that absolute and proportional represent different scores. It is possible to have a decrease in the proportion of scores, but an increase in the absolute production. With proportional numbers, small absolute differences can lead to big proportional differences. To circumvent this problem, absolute numbers are used in the analyses. In chapter 5, these absolute numbers were checked for possible influences of other, independent variables.

4.3 The children, the recordings and the transcripts

Introduction

Frequency scores of function words in Dutch child language reflect the use of function words. The goal is to test with these scores Brown’s findings in Dutch. Before I can present analyses of child language I introduce both the children, the language, and the method behind the transcription.

The children

The data of the children in this study were collected in the ‘First Language Acquisition’-project. The aim of this project was to obtain a detailed description of language development. Originally, there were 7 children in the project, but one child, a girl, acquired *otitis media*. She was left out from the analyses, and this reflects an important aspect of this study: only healthy children, without neurological, motor or other impairments (e.g. deafness), were included.

The period of observation captures a considerable portion of the early phase of language acquisition, starting at the 'single word stage' and ending at the point where the children showed an essential command of Dutch syntax. In all, this period amounts to approximately 16 to 20 months. All children were recorded in their homes. This home and normal activities situation makes it possible to obtain an optimal comparison of all children and all sessions. It is assumed that this daily routine rules out possible unnatural variation: the variation that is still in the language data is error variation. The children were engaged in their daily routine. Apart from the mothers, one other adult was present during the sessions, namely an investigator, who took part in the conversation as a normal speaker (i.e. the investigator plays with and talks to the child in the same way the mother does). The data of the children in table 4.1 have been added to the Childes Database (august 1997) as The Groningen Dutch Corpus. This corpus was collected and transcribed by Gerard Bol, Evelien Krikhaar, and Frank Wijnen (Bol, 1996, Wijnen, 1996, 1997a, 1997b).

Name	Recording age (y;m.d)	Sex	No. of recordings
Abel	1;10.30 - 2;08.13	Boy	17
Daan	1;08.21 - 3;03.30	Boy	34
Josse	2;00.07 - 2;09.16	Boy	26
Matthijs	2;00.24 - 3;07.02	Boy	42
Peter	1;05.09 - 2;08.22	Boy	27
Tomas	1;07.05 - 3;01.02	Boy	26

Table 4.1. The children in this study.

In total there are 6 children and 172 recordings. The average of age of recording at the start is 1;9 (92 weeks, Sd = 10.8), and at the end is 3;1 (159 weeks, Sd = 14), the average number of recordings per child is 28.7 recordings (Sd = 7.74). Curiously, all children are boys. This is not a deliberate choice, but purely coincidental. Whether this is or is not problematic for the interpretation of the data is discussed in chapter 6.

The recordings and the transcripts

The frequency of the recording sessions was mostly two-weekly, except for the period that the child was sick or on holidays with his parents. This exception due to sickness and holiday immediately leads to a problem. Children do not share the same weeks of recording (see table 4.1). The first recording of the children has been on a different age. Whereas, for instance, Peter's first recording was made around the 17^{th} month of age, Matthijs was first recorded around his 25^{th} month of age. The differences in the ages of the first recording were not deliberately chosen, but the result of a random selection. The only guarantee had to be that the child was still in his early stages of language development, i.e. the one word stage. The last recordings of each child also differ. Furthermore, the weeks between the recordings are, logically, not included: there are no data collected in these weeks. For example, one child is taped in week 100 and another child is taped in week 101. Therefore, it is not possible to compare children with respect to age differences and/or to average over age, which is one of the goals of this study. A solution to this problem is given in chapter 5 (using a technique called *Splines*).

The recordings of the children were transcribed according to CHILDES conventions (MacWhinney & Snow, 1990). The frequency of function words, length of utterance, mean length of utterance (MLU), and the number of words were calculated with CLAN software (MacWhinney, 1991). These frequency counts were absolute counts of tokens, i.e. the total of absolute numbers of a variable, as well as types, i.e. the sort variable.

4.4 Summary and conclusions of the chapter

A method that can reliably distinguish discontinuous change from random fluctuation or continuous change was presented in this chapter. This method consists of the bare necessities that are used to apply frequency analyses to language development in order to find evidence for quantitative change. Function words were introduced, and these words constitute a set of closed class words which are grammatical markers in a sentence, are syntactic in nature and do not refer to definite entities in reality.

There are three hypotheses with respect to the syntax of language development. The *lexical thematic* hypothesis states that all relationships between words and phrases are non-syntactical at early acquisition stages, the *reduced competence* hypothesis leaves a few blanks in the initial syntactic knowledge of a child, and the *full competence* hypothesis states that children have full knowledge of grammar structures with birth. The reduced competence hypothesis is adopted in this study.

Linguistic theories reason about the development of functional categories. If in the initial phase of combinatorial speech, grammars lack functional categories, they develop through *maturation*. An alternative hypothesis is that all functional categories and the associated grammatical modules are available from the start, but that a certain amount of input information is needed for them to become operative (*lexical learning*, learning by generalization). *Parameter setting* assumes that all linguistic knowledge is present from the beginning (a continuity approach to language development), but that it takes time to switch all linguistic buttons to their correct state. All principles may be available, and all of the necessary triggers are accessible, but that particular parameters cannot be set until certain other parameters have been set first.

Predictions of the quantitative development of function words on the basis of linguistic theories are very general. The same predictions as in chapter 2 come back again: learning theories (like Lexical Learning) would predict a gradual non-stagewise change. Change can impossibly be discontinuous. Parameter setting may explain and predict sudden changes, as all “switch” theories predict. Maturational accounts of development (in the true sense of the word maturation) are growth theories, which predict gradual change.

General observations of the development of function words can be found in the work of Roger Brown. Brown and coworkers showed that function words develop rapidly. For this reason (the sudden introduction), two stages can been distinguished: telegraphic speech, i.e. sentences have limited length, no function words, and normal, i.e. no violation of the grammar of a language, word order, and a differentiation stage.

It is assumed that all functional categories are interrelated, because they address the same syntactic slots. This interrelationship leads to the expectation that they emerge at the same time. That is, the development (i.e. the introduction) of function words is a process in which several function words develop at the same moment, and the quantitative behaviour of function words is predicted as a sudden introduction in the language of a child, possibly referring to a psychological discontinuity.

This study aims at a detailed analysis of function words. The ideal method using frequency analyses could be called ‘The 7-11 shops in America’ method. These shops are open all day, all year. With respect to development, this implies that every word on every day of every child would be recorded and analysed during a human’s life. The 7-11 method is far too labourious. With respect to syntax, a two-weekly *interval* should suffice in order to cover the developmental path. The *length* of the recording session should cover so many utterances that an increase of recording time would not change the (average) score of a certain variable any more. In this study one hour was used. A *standardized* sample means that if children are to be compared with each other, the situation in which the children have been recorded must be the same. Frequency scores are a measure of the *use* of a variable in the language of a child. These scores are indicative for functional phrases, and therefore, frequency scores refer to the acquisition of grammar, they are needed for the detection of catastrophe flags.

The subjects of study are six healthy Dutch boys with no apparent language or developmental problems. The recordings started at the 'single word stage' and ended at the point where the children showed an essential command of Dutch syntax. This period lasted roughly 16 to 20 months. All children were recorded in their homes, while the children and their mothers were engaged in their usual activities and routines. Each session lasted for approximately one hour. The frequency of sampling should be sufficient to capture the relevant fluctuations in the growth process. The data were gathered at intervals of roughly two weeks. The recordings of the children were transcribed according to CHILDES conventions (MacWhinney & Snow, 1990). Frequency counts (e.g. function words, the (mean) length of utterance, and the length of words) were performed with CLAN software (MacWhinney, 1991).

Function words are linked with important syntactic features, and they are assumed to develop rapidly. The question to be answered is:

> Under the assumption that, according to the literature, function words are introduced very rapidly (at least) and that there are two stages (a telegraphic stage and a differentiation stage), the question is if there is any evidence of a catastrophic change (as defined by the flags) in the development of function words in the six children between 1;6 and 3;0 years of age, measured on the basis of frequency analyses in a longitudinal design using two weekly observations.

In Chapter 5, I present the analyses of function words. Both continuous and discontinuous models are tested against the data.

5

Going the Distance: the results

He's going the distance
He's going for speed
...
Because he's racing and pacing and plotting the course
He's fighting and biting and riding on his horse
He's going the distance

The distance by **Cake** from their CD **Fashion nugget**

5.0 Introduction

In this chapter two approaches to development are used to find evidence for either continuity or discontinuity. The strongest claim that can be made about development is discontinuity because a greater number of conditions have to be met. I use Catastrophe theory (see also chapter 3) to find phase transitions in language development. The two sources of evidence are, first, the catastrophe flags as indicators for discontinuous behaviour, and, second, *Cuspfit* (Hartelman, 1996) that tests three models of development: a linear, a logistic and the cusp model. The data of the six children that possibly exhibit catastrophic (discontinuous) change are presented in section 5.1 (see chapter 4 for a discussion on the children and the variables).

Since the evidence in favour of a catastrophic change is not very strong, I present analyses with continuous models. Continuous models maybe linear, polynomial or logistic. The fits are based on the growth curves, with on the x-axis the independent variable (e.g. time or age), and on the y-axis the score on a variable. The important characteristic, however, is continuity in a mathematical sense. Every point in time, i.e. on the x-axis, has a score, i.e. a corresponding point on the y-axis. Section 5.2. describes how the continuity hypothesis was tested.

At the end of the chapter, the question is addressed whether the change found may be explained by a change in the input, or a change in the child's productivity (§ 5.3)

The issue at hand is the form of change. This change is either linear, non-linear, or discontinuous. The following questions are answered to choose from possible change in the development of function words:

1. What forms of change are found in language development (i.e. function words)?
2. Does the empirical evidence support discontinuous (i.e. catastrophe flags), non-linear continuous, or linear development?
3. Is there supplementary statistical evidence for either linear, continuous non-linear or discontinuous development?
4. What are the explanations that can be given for the sort of change?

Since the catastrophe flags provide the strongest (i.e. discontinuous) claim, I discuss them first. The basis of all analysis in this chapter are the dense time series analyses of the variables of all children, and they are presented next.

This chapter discusses, as the title implies, the distance a child covers. Not only in the sense of time, but also in the sense of acquiring the rules of a language (i.e. quantitative change).

Data

Quality of the variables

Before I present the quantitative analyses I want to step back from the numbers and pay some attention to the quality of the data. The data do not have an adult quality, i.e. they show various differences with mature speech: phonological, semantic and syntactic differences. In this section I discuss the first two levels (phonology and semantic), the syntactic one is discussed later (in chapter 6) in more detail.

The leaving of telegraphic speech is marked by changes on the phonological level. One of the most obvious changes is the introduction of the schwa, Ə (represented as *&6*), the 'uh'-sound in 'pres&6nt' and 'represent&6d'. In the next examples, these schwa's are used in different positions. They may be articles (5.1), particles and expletives like *er* 'there'(5.2), but in many cases, it is unclear where they stand for (e.g. it could be a personal pronoun) (5.3). There are two translations: one literary and one in anglicised, 'normal' English (in *italics* and with possible necessary additions in [brackets]).

5.1. a. 'dat &6 vlinder.' (Abel, age 2;4)
That the/a butterfly. *That [is] the/a butterfly*.
b. 'ik nog &6 huisje.' (Daan, age 2;5)
I still a house. *I still [want/have] a house*.
c. 'Josse ook &6 suiker.' (Josse, age 2;5)
Josse also the sugar. *Josse [has / wants] also the sugar*.

5.2. a. ‘deez zit &6 op.’ (Abel, age 2;4)
This-one sits there on. *This one is on that.*
b. ‘eef &6 koud’ (Josse, age 2;5)
Has it cold. *[She/he/I] has it cold.*

5.3. a. ‘en &6 lopen kan.’ (Matthijs, age 2;5)
And &6 walk can. *And [I?] can walk*
b. ‘die heeft ook &6 geen wielen.’ (Tomas, age 2;5)
That-one has also &6 no wheels. *That one has also no wheels*.

A second aspect of early child language is that not all sounds are realised correctly. Most notably is the absence of voiceless consonants. For example, children produce *t* instead of *d* or *s* instead of *z*. Some examples[1] are in example 5.4.

5.4 a. ‘&6 tan [= gedaan]’ (Daan, age 2;4)
Done.
b. ‘de pus’ [= puzzel].’ (Matthijs, age 2;3)
The puzzle.

Third, a child in the early stages of language development fails to produce in many cases (but not all!) consonant clusters. Strings like *gr* (as in *groot* ‘big’), *st* (as in *stukje* ‘a bit’) or *kl* (as in *kletsen* ‘chat’) are mostly absent in early child language. This, however, is not a problem for the analysis of function words since in Dutch nearly all function words are free from consonant clusters.

5.5 a. ‘die goot [= grote] wielen moet ie hebben.’ (Tomas, age 2;5)
Those big wheels must he have. *He must have those big wheels*.
b. ‘tukje [stukje] vliegen’ (Daan, age 2;4)
A bit fly. *To fly a bit.*
c. ‘Ieke [= Frederieke] &kEs6 [= kletsen]’ (Matthijs, age 2;5)
Frederieke chat-inf.

Semantically, the variables themselves constitute a mixed group (see the Appendix for an overview of the function words analysed). First, there are three articles, which includes phonological and contracted forms. Phonological forms of articles are an approximation of the adult forms. Children say, for example, *&6* instead of *een* or *de* to fill a syntactic slot that is reserved for a determiner. Contracted (or clitised) forms are *‘t alarm gaat af*,

[1] The transcripts lack a thorough phonological description, and therefore an analysis of voiceless consonants was not possible.

where *'t* is the contracted form of *het*, the neuter form of 'the' in Dutch. Children start with mainly the phonological forms, that are slowly replaced by the adult forms. In other words, the phonological forms of the articles precede the correct forms, and the contracted forms are later than the correct forms. Second, there are nine modals (both finite and infinite). The modals produced are mainly finite, present and singular forms. That is, the increase of modal verbs in the early stages is not due to plural forms or / and the non-present tense of the inflected verbs. Third, there are 3 kinds of pronouns (4 determinative, 17 personal, and 8 possessive pronouns). In line with the modal verbs, the pronouns found are mainly in the nominative case and singular (i.e. *I, you, he, she* and *it*). Over 87 % of all pronouns produced by the children is in the singular person and in the nominative case. This indicates that the development of these function words has not ended yet, because there are hardly any plural pronouns (e.g. *wij* (we)) and dative and accusative forms (e.g. *mij* (me), *jou* (you - dative) or *hen* (them)) in early child language.

The diversity of personal pronouns and modals increases around the first major increase in function word production. That is, apart from an increase in the number of **tokens** there is an increase in the number of **types**. The average number of personal pronouns before the sudden increase is 1, while after the jump the average number is 9. Likewise, the average number of personal modals before the sudden increase is 2, while in the later stages of development the average number is 9. A similar increase in types in the third year of life has been found in prepositions (Sinha, Thorseng, Hayashi & Plunkett, 1994).

The number of types of articles is so small that the diversity can only increase from 0 to 3. However, what is extremely interesting to see is that the form of articles is not confined to the adult forms *de*, *het* and *een* ('the' and 'a(n)'). The child shows a certain knowledge of determiners, which is expressed by the phonological forms of articles. A meta analysis of the data reveals that articles, and determinative and possessive pronouns occupy the same position in a sentence, i.e. they are all considered as instantiations of the syntactic category Det. All three appear before a noun. According to syntactic theories, this means that these three classes of function words appear in the same syntactic slots.

From the qualitative analysis (a syntactic discussion will follow in chapter 6) it follows that there is quite some diversity in the language in a child. **The** question from here on is: do the data support a discontinuous or a continuous hypothesis? Before I test both hypotheses, I present the individual growth curves.

Individual curves

The graphs in Figure 5.1 below contain the raw scores of the three classes of function words (articles, the three categories of pronouns and modal verbs). That is, the absolute number of function words is measured during each session, for each child individually. These numbers, i.e. frequency scores, are plotted as a function of age.

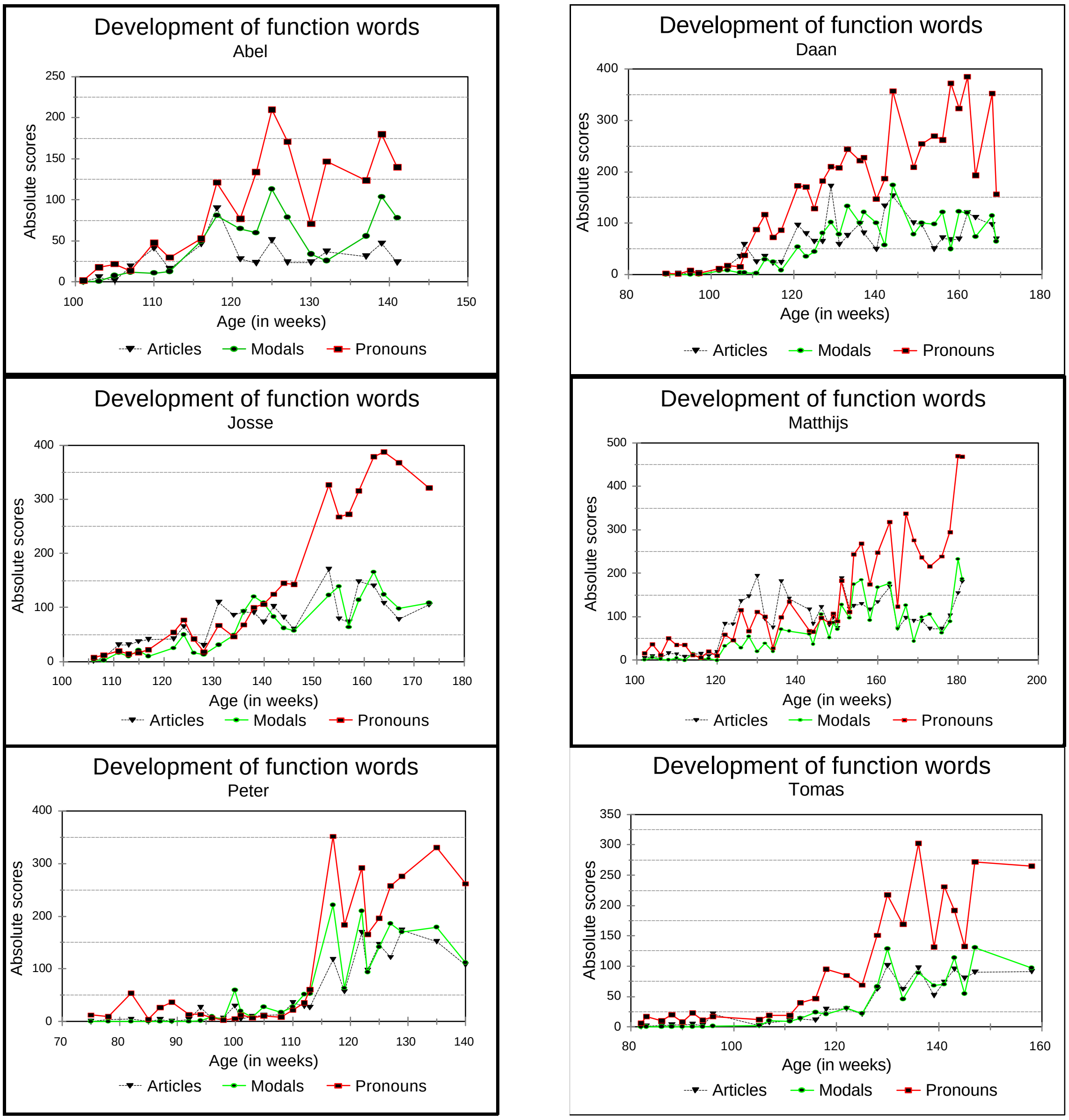


Figure 5.1. The development of function words: all recordings.

At first sight, the data suggest a number of general characteristics. First, some children seem to show a sudden increase (e.g. Peter), whilst some take it easy and gradually increase their production of function words (although there are oscillations in the amount of function words, e.g. Abel). Second, the exact moment of the increase (i.e. the age of jump onset) differs considerably amongst children. For example, Peter starts at the 115th week, Josse starts at the 145th week. This means that age is probably not a good predictor of the onset of the jump, although a rough prediction like 'between the age of two and three' would eventually suffice (see also the subsection of the continuity hypothesis for a discussion on age averages). Third, there are differences between the children with respect to the absolute number of instances of function words. For example, Abel has an average number of pronouns between 150 and 200, Daan is between 150 and 400. Furthermore, there are differences between the variables. For example, in all children the amount of pronouns is bigger than modals and articles. Within a child, the start of the use of the three main categories articles, modal verbs and pronouns takes place in the same age range.

In sum, there are large individual differences and differences between variables. The question is: given these individual differences, is there any evidence for discontinuous change in terms of the catastrophe flags? I will analyse the data for each flag that may be found in longitudinal data separately.

The catastrophe flags: indicators of discontinuity

As I mentioned above, not all catastrophe flags can be found in spontaneous and longitudinally recorded language development, since some flags demand a perturbation of the system (see section 2 in chapter 3). The analyses of four out of five possible flags in the longitudinal data are presented here.

Sudden jump

The first flag refers to jumps in a variable (i.e. in a time series). A sudden jump is an increase in a score. It can be defined as the growth between two sessions which is bigger than any other growth increases between any two other sessions, but since the empirically observed data show the performance level of a variable, not its abstract or potential equilibrium level, a problem occurs with distinguishing a mere increase in growth rate from a sudden jump (a discontinuity) between two equilibrium modes. To test the jump, the data were fitted with a logistic equation.

The fitting procedure

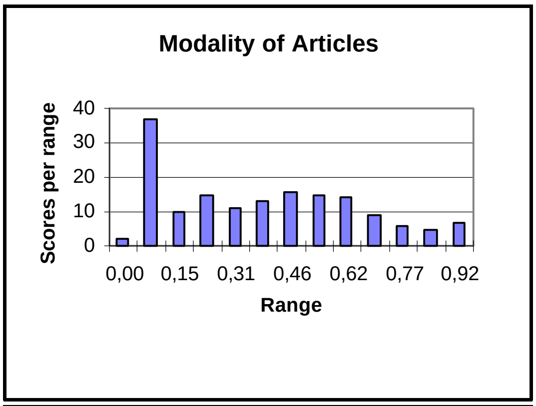
Magestic contains a fitting procedure in which any model may be postulated. It is a program that uses the mathematical routines of a spreadsheet application. The first step is the definition of a model, for example, a linear or a logistic model. *Magestic* will look for combinations of parameters in the specified model. For example, in a linear model, defined as $y = \alpha + \beta.x$, the program estimates α and β. If values of parameters, for example α or β, may not exceed a fixed number, boundaries for the parameters may be defined. In order to find a jump, the logistic model (see also chapter 3) was fitted. The assumption was tested that these curves represent jumps in the frequencies by estimating their parameters. The exponents and growth rates indicate whether the curves are jumps or not. If the exponent approaches 2 and/or the growth rate approaches 1, the resulting growth curve takes the form of a steep cliff-like change, i.e. a jump (see van Geert, 1994). In the next table, the values for the parameters are given (Ruhland & van Geert, 1997).

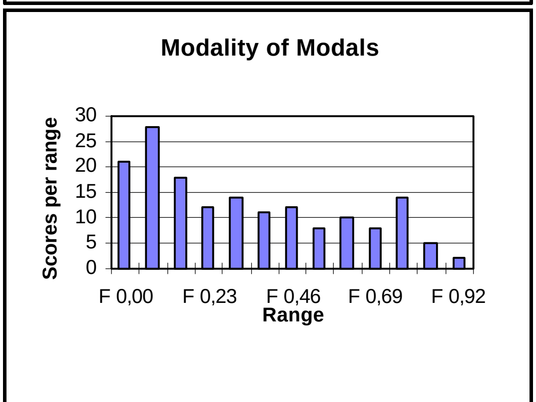
Pronouns	Abel	Daan	Josse	Matthijs	Peter	Tomas	Average
initial level	0	0	0	0.049	0	0.029	0.024
exponent	**1.46**	0.47	**1.61**	1.092	**1.49**	**2.012**	**1.356**
rate	**2.74**	0.12	0.68	0.1	**3**	**2.948**	**1.6**
car. capacity	0.66	0.77	0.93	*2*	0.73	0.696	0.748
Modals							
initial level	0	0	0	0.046	0	0.01	0.017
exponent	1.28	**1.5**	0.92	**1.932**	1.12	**1.549**	**1.382**
rate	**2.1**	1.23	0.29	1.002	**2.05**	**2.233**	**1.484**
car. capacity	0.62	0.58	0.7	0.55	0.68	0.671	0.633
Articles							
initial level	0.1	0.1	0.1	0.013	0	0.011	0.035
exponent	**2.04**	**1.76**	0.48	**1.534**	**1.94**	**1.41**	**1.527**
rate	***5***	1.21	0.13	**2.158**	**2.51**	0.822	**1.971**
car. capacity	0.45	0.53	0.67	0.614	0.79	0.689	0.624
total averaged							
initial level	0	0	0	0.036	0	0.015	0.025
exponent	**1.59**	1.24	1	**1.519**	**1.52**	**1.657**	**1.421**
rate	**1.61**	0.86	0.37	1.086	**2.52**	**2.001**	**1.685**
car. capacity	0.58	0.63	0.77	1.054	0.73	0.685	0.669

Table 5.1. Growth parameter estimations based on the individual data series.

The figures in bold (the grey cells in the table) refer to steep curves and could be indicative of sudden jumps. Out of 18 cases (6 children x 3 variables), 2/3 of the parameter values hint at a sudden jump: either the exponent p is over 1.5, or the rate, r, is over 1. As we discussed in Ruhland and van Geert (1997), we used this as a demarcation criterion, since there are no absolute values of these parameters that can be mapped on the data directly.

Bimodality





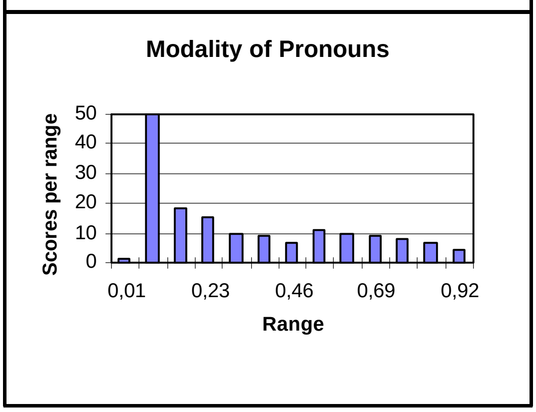


Figure 5.2 The modality in the data

The analysis of bimodality is in figure 5.2. The data were rescaled to a 0 - 1 range (remember that the form of the curves is important, not the actual productional level). All the scores of each variable separately of all the children are used to determine bimodality. A step-range was determined by slicing the y-axis into, for example, 10 equal steps, and counting the number of scores in each slice. This step-range, the slice, is on the x-axis, on the y-axis is the number of scores in each slice. In other words, I performed a frequency count of the scores *of a session* that fall in each range/slice (see also Wimmers, 1996). This leads to a histogram that consists of the frequency count of scores in 12 steps (Figure 5.2).

The graphs show more than one 'bump', but the question is whether these bumps reflect unimodality or multimodality. The first criterion for a transition is the presence of a bimodal frequency distribution in the scores (see above). To determine the existence of two bumps, bimodality is tested statistically by a mixtures distributions analyses, and with a kernel density method.

Since the histograms show several peaks in the slices, the scores in those slices apparently do not show a standard unimodal distribution, i.e. not one peak. The question is if these peaks represent multimodality. A statistical test for multimodality in the slices is available, namely a mixture distributions. This mixtures distribution analyses determines the number of modalities (Everitt & Hand, 1981; see also van der Maas, 1993 and Wimmers, 1996 for an application). The analysis is a k component mixture which is defined as:

$$f(x;\theta_0,\pi_0,n) = \sum_{()=1}^{k} \pi_0 b(x;\theta_0,n)$$

with θ the chance of success per component, π the proportion of scores per component, and n the number of trials. The variance of the equation is

$$\sigma^2 = \sum_{()=1}^{k} \pi_0[n\theta_0(1-\theta_0)+(n\theta_0)^2]-\mu^2$$

and the mean is

$$\mu = \sum_{()=1}^{k} n\pi_0\theta_0$$

Depending on the number of components (e.g. for k = 1 the model equals a model with 1 equilibrium), k values of θ and k - 1 values of π have to be estimated. θ is the chance of success per component and π is the proportion of scores per component. Both θ and π are distribution parameters.

The best fitting mixture model (e.g. a distribution model with 1, 2, 3 or even more modes) was chosen on the grounds of three selection criteria. First, we used Akaike's Information Criterion (AIC). AIC is defined as minus twice the loglikelyhood plus twice the number of parameters. This value has an extra weight, based on the amount of parameters, and the model with smallest AIC is the best (see also the section on the Cuspfit). AIC 'punishes' for the number of parameters in a model, or in other words, it compensates for the number of parameters. The second criterion is χ^2, the goodness of fit between the empirical data and the data from the models. The smaller its value, the better the model fits the data. The third criterion, VAF (that is, variance accounted for under any model) is obtained through a ratio of the variance of the binomial in which the model parameters are replaced by model estimates and the sample variance (Wimmers, 1996). A large VAF indicates a satisfactory model for the data.

Model	θ	π	AIC	χ^2	VAF
1 component	0.39		1125	6893	0.26
2 components	.12 .61	.46 .54	852	130.2	0.86
3 components	.08 .92 .47	.35 .13 .52	797	4.21	0.99

Table 5.2. Model estimates of *articles*.

χ^2 degrees of freedom are 9, 7 and 5 for 1, 2 and 3 component model.

Model	θ	π	AIC	χ^2	VAF
1 component	0.33		1140	19728	0.24
2 components	.11 .63	.58 .42	804	82.86	0.89
3 components	.35 .76 .05	.37 .24 .39	769	13.72	0.98

Table 5.3. Model estimates of *modals*.

Model	θ	π	AIC	χ^2	VAF
1 component	35		1198	2070815	0.23
2 components	.12 .67	.59 .41	811	601	0.91
3 components	.55 .10 .92	.36 .54 .11	776	217	0.99

Table 5.4. Model estimates of *pronouns*.

The results from the one-, two- and three-component mixtures distributions are based on a binomial distribution. New techniques have yet to be developed to test the modality of data based on distributions other than binomial ones. Since these are not available, I chose binomial. This means that the results from these analyses (tables 5.2, 5.3 and 5.4) have to be interpreted with caution.

The analyses are in favour of a multimodal distribution. In tables 5.2, 5.3 and 5.4 the results of AIC, χ^2 and VAF are similar: AIC and χ^2 decrease, while VAF increases. A 1-component model does not fit the data well. The difference between the 2 and 3 component model, however, is small. It cannot be concluded that there is a unimodal distribution.

In order to check whether the three-modal result is not an artefact of the statistical procedure chosen, we also tested for modality with the aid of a more robust procedure, a kernel density technique. The application of a mixtures analyses and the kernel density technique for testing the number of modes (Silverman, 1983) are different techniques. The mixtures method tests the number of classes. Each class has its own distribution of the behavioural variable, that is often normal distributed. The overall distribution is obtained by averaging the distributions of each group. Mixtures are not very reliable since a two classes do not implicitly lead to a bimodal distribution if there is large variance. It would seem that there is a unimodal distribution. Silverman's algorithm uses a kernel density estimation and bootstrapping (see also Silverman, 1983). It cannot be used to detect the underlying classes (as in the case of a mixtures distribution), but it has the advantage of being non-parametric.

The kernel density method (Silverman 1981, 1983) has a nasty drawback. If the data are analysed with less modes than there presumably are in the data, the p-values become unstable. The data do not show a monotonous pattern. That is if the analyses show the following monotonous pattern, it is quite clear that there is a probability that there are three modes.

Modes	1	2	3	4	5
p-value	0.1	0.15	0.85	0.9	1

However, if they show the following pattern, non-monotonous pattern, there is no clear-cut interpretation possible.

Modes	1	2	3	4	5
p-value	0.1	0.15	0.85	0.6	0.7

The data in the next table show that there is not a monotonous pattern, and this makes an interpretation of the data problematic.

	Modes		
Variable	1	2	3
Articles	0	0.1	0.81
Modals	0.6	1	0.66
Pronouns	0.3	0.4	0.63

Table 5.5. Output of the kernel density technique (with simulation = 100, p-values = 3)

The probability that a certain variable has one or less, two or less or three or less modes is not clear for the modals and the pronouns. Only in the case of the articles there is a high probability that there are three modes: the probability is .04 that the number of modes is 1 or less (i.e. 1 or 0), a probability of .10 that there are two or less modes, and a probability of .81 that there are three or less modes. I used the relative (standardized) scores of the three variables. Strangely enough, there are large differences between the outcomes of the analyses of both the relative and the absolute data. This may hint at a problem with Silverman's method. Finally, I tested the three variables for more than 3 modes but with any high number of grid steps and any high number of simulation, the chance is very small that there are more than three modes.

Inaccessibility

It is assumed that there is a region that is *inaccessible*: behavioural variables cannot be in a state somewhere in between these two possible behaviours. This flag prevents the conclusion that a gradual change, in which case there are also two states, is a case of qualitative change. Inherent to psychological research, it is accepted that there is a probability that scores fall in the inaccessible region (see also Wimmers, 1996). Evidence for this flag means that there is a lack of scores in between the two equilibria (or modes). That is, scores do not fall in between the two modes, the system cannot access this region. A look at the time series of Figure 5.1 reveals that if there are two equilibria, and we shall assume that this is the case, the change from one to the other equilibrium is so sudden that there are no scores between the two equilibria. The flag inaccessibility in the frequency plots (Figure 5.2) should have been between the peaks. As can be seen in Figure 5.2, there are many scores in the assumed inaccessible region, but this is not as strange at it seems, since there is, in accordance with Wimmers (1996, 61), a *chance* that a score will fall in the inaccessible region. Furthermore, it might be the case that the language of some children does not change suddenly (i.e. in catastrophic terms), because the recordings do not cover the complete age span of function word development. In the next chapter I argue that this is probably the case in Abel’s development.

Anomalous variance

The data are not very appropriate to demonstrate the flag *Anomalous variance* (AV), since the frequency scores in figure 5.1 are additive scores. If there had been a plot of all tokens for every small time unit (e.g. a second) during each session, AV could have been found in a session by using a moving window, since such a time window would show the variance during the smallest time unit, and therefore show the sensitivity of a child to small changes. Unfortunately, such a plot is not available. Using “variance within sessions” as an index of anomalous variance might be an incorrect index. The transition period is a period in which the language of a child is changing within weeks (or even days). Using a variance within sessions does not refer to this transitional change over weeks. Therefore I used the method explained below.

Normally, it is expected that variance increases with age, under the assumption that the productivity or level of a variable increases. An increase of scores of a variable leads to more variance, i.e. the chance of large variance at high values is bigger than the chance of large variance at low values. However, one of the flags, AV, assumes that there is no continuous or linear increase of variance around a catastrophic transition.

That is, AV indicates that there is a large increase in the variance around the sudden jump, which is not concordant with the (linear) variance which would be expected.

Anomalous variance was translated to observational values. The idea is that around the catastrophe the variance increases. There are several ways of applying this flag to the data. First, there could be an increase in variance *between* the sessions. This variance is expressed in fluctuations over time (e.g. age). Second, there might be an increase in variance *within* the sessions. That is, in the case of language development and under normal circumstances, assuming that the system is stable, it is expected that a child is constant in his production of language. However, in the case of a transition, this anomalous variance flag should indicate that, with respect to our longitudinal data, a child is sensitive to all sorts of influences, internal and external. The instability which is caused by these influences is reflected in the amount and variability of speech production during one hour of recording. Normally, one would expect approximately the same amount of language production in, for example, the first and the second half of one hour of recording. If Anomalous variance applies, this expectation is probably not fulfilled. So, I chopped the recordings of one hour in exactly two parts and counted the number of occurrences of a variable during each half hour. The assumption is that in a transition period function words are not distributed evenly across all sentences recorded during an observation session, but that they occur in spikes or bursts. That is, if the use of function words (e.g. pronouns) corresponds with an unstable state of the language production system, it is likely that function words are used in a highly irregular way (see for instance Wimmers, 1996, for a comparable approach to the transition in reaching and grasping). Sometimes the child behaves as if he masters the use of pronouns, whereas at another occasion it looks as if he is still in his pronoun-less stage. It is expected that, if a particular grammatical phenomenon, such as the use of function words, is a novelty, events such as inviting the child to start playing in the presence of a relative stranger (the observer), setting up the observation equipment, and so on, will in fact act like a mild perturbation. If the language production system is in an unstable state with regard to the grammatical phenomenon at issue (e.g. the proficient use of pronouns), mild perturbations cause a divergence of linear response, or a critical slowing down. It will take some time before the system has reinstalled its equilibrium, that is, the stability required for a productive use of, for instance, pronouns. From this assumption follows that during the transition stage there exists a high probability that the frequency of function words, pronouns for instance, is considerably higher in the second than in the first half of the observation session. In order to test this assumption, the pronoun frequency[1] in the second half was divided by that of the first half. In order to compensate for the fact that observations with very low frequencies, which were obtained in the first phase of pronoun development, eventually show high, merely accidental fluctuations between the halves, a constant was added to the frequency counts. This constant had the size of one standard deviation of the individual time series of each child. Second, I reasoned that considerable fluctuations between the halves have more weight if they occur with highly productive use of pronouns than if they occur with very low pronoun frequencies (that is, the ratio 200/100 is a considerably more significant fluctuation than the ratio 2/1). Multiplying the ratio by the total frequencies of function words during the observation session at issue would however, distort the ratios too much. Instead, the ratios were multiplied by the logarithm of the total frequency during the observation session. Thus, if f_1 is the frequency of pronouns used during the first half of an observation session n, and f_2 is the frequency during the second half, the observation session's halves ratio H_n is computed in the following way

$$H_n = \left((f_2 + c) / (f_1 + c) \right) . \log (f_1 + f_2)$$

for H_n = halves ratio, c = constant (1 standard deviation), f_1, f_2 = frequencies in first/second half of session.

I plotted the H ratio's for every child for the variable pronouns and compared the maxima of the ratios with the corresponding total frequencies. It is expected that, if anomalous variance holds during the transition, sudden jumps in the use of pronouns should coincide with strong deviations in the halves ratios (assuming that, on average, the frequencies of the first and second halves are about equal). Visual inspection of the plots of the pronoun frequencies and halves ratios indicates that this coincidence occurs with two of the six children, namely Peter and Tomas (see figure 5.3). The data from Peter and Tomas show a clear jump-like increase in pronouns. Josse, however, whose pronouns also showed a significant jump, did not show the expected anomaly in the frequency halves during the jump. In summary, the evidence, based on pronouns, for anomalous variance that occurs during the jump is rather weak.

Not all children show anomalous variance (i.e. with a visual inspection). Contrary to Peter's and Tomas' case, I did not find a difference between the first and second half of the recording session of, for instance, Abel's function words development. This could mean two things.

[1] I present only one example (pronouns) to prevent an overkill of examples.

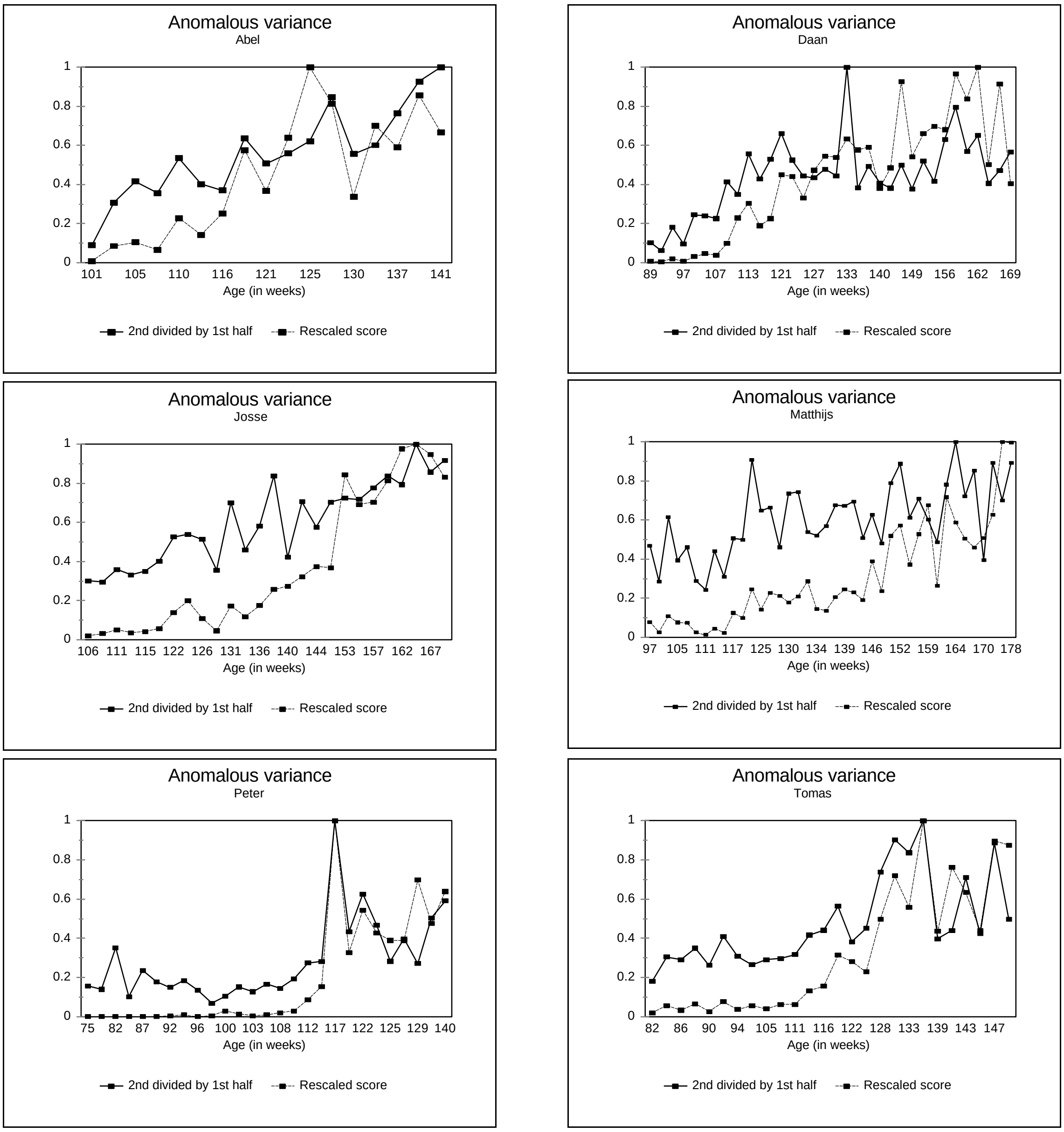


Figure 5.3. Analyses of Anomalous variance in pronouns

First, there is no anomalous variance in Abel's development of pronouns. This would indicate that a discontinuous model of language development might (as defined by catastrophe theory) not be the right model, at least for some of the children (e.g. Abel). Although other children seem to develop in a discontinuous manner, like in Peter's case, these individual differences between children are by no means undesirable, it just means that children have the opportunity to walk and talk along various paths of development (see chapters 4 and 6 for a discussion on nominal and pronominal style).

There are other ways of operationalizing anomalous variance (apart from "variance *within* sessions"). First, anomalous variance can be operationalised as fluctuations around the transition. This is variance *between* sessions. In that case, anomalous variance is a *developmental* characteristic. This type of variance can be seen in nearly all the children and all the variables. However, it is unclear how these fluctuations should be tested statistically. Second, a comparison can be made between absolute and relative scores of function words. *Relative* means that the behavioural variable is dependent on some other independent variable (other than age), and a peak in this ratio would indicate anomalous variance. The increase of function words could be directly related to the increase of talkativeness. That is, if articles, for instance, constitute a certain percentage of language (as a baseline), than there is a gradual increase when language productivity increases. However, if children use extra articles around the jump (as if they are 'playing' with articles), then there is an increased percentage of articles that will disappear after the jump. In figure 5.7 (section 5.2) I present such an analysis of the behavioural variables as a ratio of vocabulary (words spoken during a session). There is no peak in the ratio, and therefore, there is no indication of anomalous variance. The variance either slowly increases, or stays constant over time. Apparently, this ratio does not shed any new light on the anomalous variance flag.

Direct Catastrophe modelling: Cuspfit

Sussman and Zahler (1978) criticised the lack of a statistical underpinning of catastrophe theory. They make an important point, because the flags are nothing more than indicators of discontinuous change. The flags do not provide strong evidence from a statistical point of view, they merely suggest that the change found in six children is a catastrophic change. Nearly twenty years later, much work has been done to counter this criticism. It is now possible to apply more rigorous statistical testing with the aid of the statistical program developed by Cobb (Cobb, 1978, 1980). Recently, this program has been improved and expanded by Hartelman (1996). Hartelman's program, *Cuspfit*, has two drawbacks with respect to this study. First, it takes as an input cross-sectional data. The data in this study are longitudinal, but I nevertheless decided to treat them as if they were cross-sectional. Second, it is necessary to have a control parameter for the cusp to occur. Because these parameters are unknown, I treated age as the independent variable under the assumption that control parameters are related to age. The output of the program is, first, a bifurcation diagram, showing how the scores fit into the bifurcated region (see also Figure 3.1, the cusp). The scores of the variables (here: articles, modals and pronouns) and the corresponding age (which is nothing more than an index of time, and not the decisive independent variable) are tested against three models, namely a linear, a logistic and a cusp model. The results of fitting of all the data are in tables 5.6 (i.e. the three models are tested against the empirical data of articles, modals, and pronouns, and the output of the testing yields a pseudo-r^2, the explained variance).

Model	Articles (r^2)	Modals (r^2)	Pronouns (r^2)
Linear	0.4328	0.4519	0.4883
Logistic	0.5362	0.5219	0.5113
Cusp	0.6541	0.3408	0.7054

Table 5.6. Output of Cuspfit: pseudo-r^2.

The best fit occurs in the case of the maximum amount of variance explained. However, it is not possible to test the various models against each other, since each model is not a special case of the other two models. That is, the models are no simplification of each other, there is no ‘nesting’ of the models. In two cases (that of articles and of pronouns), the cusp model yields the best fit, but with respect to the fit of modals, the best fit comes from the logistic model.

These pseudo-r^2 values do not take into account the number of parameters of the three models. This is important since an increase of parameters will automatically lead to better fits. Fortunately, Cuspfit also renders the AIC (Akaike’s Information Criterion) value. In table 5.7, these values are listed, and here, the lowest scores indicate the best fit.

		AIC	
Model	Articles	Modals	Pronouns
Linear	618.88	625.05	637.44
Logistic	382.54	387.99	391.92
Cusp	306.87	302.42	289.69

Table 5.7. Output of Cuspfit: AIC.

The AIC gives the best fit more plausibility. Since there is no nesting of the three models, the differences between the AIC cannot be tested statistically. In all cases, the cusp model receives the lowest scores, and therefore, the data are fitted best by the cusp model (see for more details on the technical aspect of the output of the program, Hartelman, 1996). In all, the output of the program hints at the cusp model as the best model for the data. There has to be some caution, though. Suppose the sampling of the data is composed of s-shaped curves. This may lead to a best fit of the cusp, since the sum of these curves lead to a higher density of the initial and the end stage of the summed curves. Figure 5.4 shows this quite nicely: the two dark areas (around the 0 score and the 90 score) have a higher density (more points from the curves) than the light grey middle that is an area with a low density. As a result the cusp will be the best fitting model, although the individual curves are all s-shaped (e.g. the s-shaped curves of Wells, 1986).

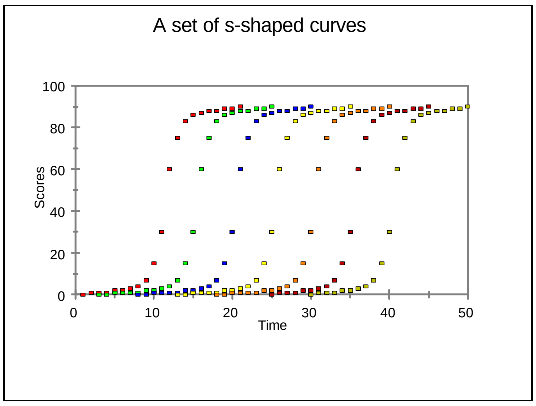


Figure 5.4. A set of s-shaped curves.

Summary

In sum, there is reason to believe that something more is going on than the standard linear continuous assumption of development (i.e. the development of function words), because there is some evidence for discontinuous change, namely the occurrence of a number of catastrophe flags. A *sudden jump* was found in all three variables, but not in all children (i.e. in 14 of 18 cases; see also table 5.8.a). This flag indicates that at least the variables (at least) change very fast from one equilibrium to the next. Within two or three weeks, the production of function words reaches a new level or equilibrium. *Multimodality* was found which indicates that there is more than one equilibrium, and there are three clues for this multimodal distribution. First, the frequency distribution in Figure 5.3. shows multimodality. Second, the Kernel density technique renders complementary evidence. The results indicate that there is more than one mode, namely three. Third, the mixture distributions does not contradict the non-unimodal data of the kernel density technique (although the binomial analyses must be treated with precaution). The third flag, which should come with bimodality in the case of a catastrophic change, is *inaccessibility*, but this flag could not be found. There are too many scores between the modes (2 or more), and the system changes rapidly, but not instantaneously from one mode to the other. Fourth, *anomalous variance* was found in three of the six children for pronouns. This indicates that if there is a transition, i.e. discontinuous change, the variance increases temporarily, i.e. in an anomalous way. Unfortunately, the other variables did not show any anomalous variance.

Supplementary evidence for discontinuous change comes from *Cuspfit*. The analyses show that in all cases (articles, modals and pronouns), the cusp model is best model. A comparison of three models (linear, logistic, and cusp) shows that the cusp model is the best model to fit the data.

The conclusion must be that the development of function words in quantitative terms appears to be a sudden development. The fits with the non-linear equations indicate that the change in the individual data is jumpwise since the power of the equation is close to 2. There is some evidence for four flags, the analyses with the aid of Cuspfit suggests additional evidence for a catastrophic change (i.e. the cusp model), and there is more than one mode (or equilibrium) in the data of children according to analyses with the kernel density technique and the mixtures distribution.

In the next tables I indicated for each of the four flags whether or not the flag was found in the data (a + means found, a - means not found).

	Articles	Modals	Pronouns
Abel	+	+	+
Daan	+	+	-
Josse	-	-	+
Matthijs	+	+	-
Peter	+	+	+
Tomas	+	+	+

Table 5.8.a Sudden jump.

	Articles	Modals	Pronouns
All children	+	+	+

Table 5.8.b Multimodality.

	Articles	Modals	Pronouns
Abel	-	-	-
Daan	-	-	-
Josse	-	-	-
Matthijs	-	-	-
Peter	-	-	-
Tomas	-	-	-

Table 5.8.c. Inaccessibility.

	Articles	Modals	Pronouns
Abel	-	-	-
Daan	-	-	-
Josse	-	-	+
Matthijs	-	-	+
Peter	+	+	+
Tomas	-	+	+

Table 5.8.d. Anomalous variance

The question is: is the change found a discontinuous change, as defined by catastrophe theory? More specific: can we conclude that the change of function words in language development is best fitted by a cusp catastrophe? The answer must be a provisional no. The evidence from the flags and the Cuspfit does not apply to all children, an overall conclusion of catastrophic change must be rejected on the basis of the time series analyses. Thus, the results are not conclusive, since not all children show, for example, a sudden jump or anomalous variance. It must be said again that not all Catastrophe flags can be found in longitudinal data.

In sum, there exists large individual variability in the development of function words in each child. The literature suggests that the development of function words is a sudden development. Using dense time series I showed that there is indeed a sudden jump in most of the children. However, there is insufficient evidence to support the hypothesis of discontinuous change as defined by Catastrophe Theory and simultaneously reject the continuous change model. The next step in the analysis is a test of the continuity hypothesis using continuous models to see what they explain regarding the data, since the discontinuous analysis is not conclusive.

5.2 Continuous models of development: evidence from the data?

Introduction

In the previous section, the discontinuity hypothesis was tested. However, discontinuity could not be demonstrated sufficiently for all children. In this section, I will search evidence using models and statistical techniques to test the continuity hypothesis.

The question is whether the evidence for continuity is linear or non-linear. The continuity hypothesis states that the paths found in the individual data are either linear or non-linear, but in any case they are continuous. That is, continuous models are models in which each independent variable (here: age) has a score on the dependent variable. Continuity means that between every two scores or levels there are intermediate points or levels. The question is whether linear models yield better fits than non-linear and continuous models (see chapter 3 for a discussion on continuous models).

Method

There is a nearly infinite number of continuous models, but from a developmental perspective, only a limited number of these models make sense. A higher order polynomial will lead, especially if the order of the equation is equal to or higher than the number of data points, to a perfect fit of the empirical data. This is undesirable since a higher order needs to be linked with psychological parameters, and this does not seem realistic. So, the models that have been chosen here are the simplest models, i.e. a linear model and the growth model.

Continuous models: linear and non-linear fitting

I fitted two models against the data of the children: a linear model and a logistic model. The testing of continuous models (either linear or non-linear) is done by using simple tests like standard deviation, variance and regression. The intention of fitting the data is to find a model (i.e. an equation) for the developmental path found. Both the growth model and the linear model (see chapter 3) are fitted with the aid of *Magestic (*see *The fitting procedure* in section 5.1). The outcome of *Magestic*, i.e. the best fit, is defined by the least sum of squares.. The parameters (initial value, growth parameter, etc.) that control the equation are estimated accordingly. In Table 5.9, the least squares of sums are given, not the explained variance, since a linear fit will also lead to a high r^2 if the line runs through the middle of the data points. It is therefore not a good index for the quality of the fit.

Variable / Model	Abel	Daan	Josse	Matthijs	Peter	Tomas	Average
Articles *Linear*	2.696	6.822	4.407	8.507	6.346	5.077	5.643
Logistic	2.335	4.124	3.822	5.088	2.339	2.171	3.313
Modals *Linear*	4.248	7.617	6.137	7.277	6.3	6.03	6.268
Logistic	2.15	3.104	2.973	4.809	2.584	2.565	3.031
Pronouns *Linear*	4.088	6.546	7.626	6.847	6.478	5.781	6.228
Logistic	1.964	3.1	1.316	3.576	2.345	1.931	2.372

Table 5.9. The least sums of squares values for the growth model and the linear model respectively (Including an average of articles, modals and pronouns).

In order to compare the quality of the model fits, least sums of squares measures have been computed for the fits based on the non-linear growth model and fits based on the linear model. The average sums show that the growth models (average of sums is 2.9) are more than twice as good as the linear model (average of sums is 6.0). On the basis of the complete series of the individual data, a nonlinear growth model fits the data best. The values of power and rate are in most cases over 1 (see table 5.1).

Individual processes vs. group averages

Introduction

In developmental psychology, most research is based on analyses of group means, calculated over age. Wells (1986), for instance, used age averages in his study of the development of function words. If the data of the six children were averaged over age (i.e. the number of, for example, pronouns was counted at age t, t + 1, etc.), the x-axis is an index of chronological age in weeks. However, since the ages of observation (in terms of weeks) were not the same for all children (e.g. one child was observed on week 100 and 102, whereas another child was observed on week 101 and 103), a technique called Splines was applied to the data to handle this problem.

Splines: a technique for missing data

The recordings do not cover every week (or even every day) of the overall period of the recordings. This is problematic since one of the objectives is to average over age. The

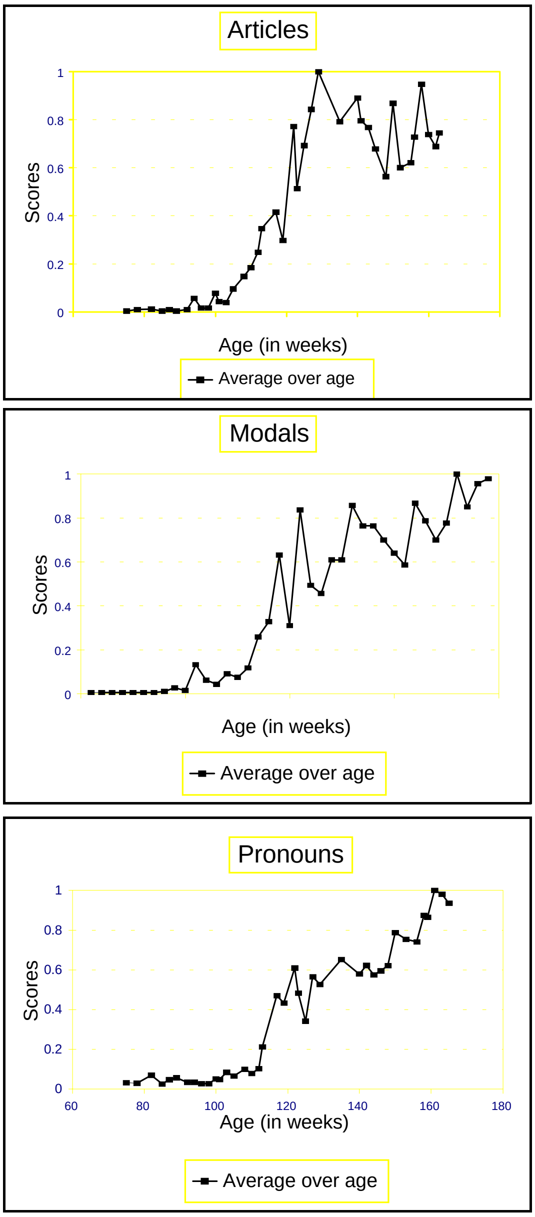


Figure 5.4. Average over age.

problem of missing data can be solved by using a method called splines. According to Hyams (1996), there are four types of spline interpolation (see the Help file in

CurveExpert). The choice of splines will have an influence on the “splined” data: estimated points will be between two data points, but they differ with respect to their exact place in the splined data and they depend on the spline that is used (i.e. linear, quadratic, etc.). Since the overall pattern of the change in language development is non-linear, i.e. sudden, the cubic spline was chosen. It is not assumed that the pattern of change in the weeks when no recordings were made follows shows a linear trend, i.e. there is no simple connect-the-dot principle. In other words, it is assumed that the spline reflects the general pattern of development, i.e. a non-linear pattern. The splines of the data result in a week-to-week data set. From then on, it is possible to average over age. This is done in *The results of age averages*.

The results of age averages

Note that, since the ages of the observations differed across the children, the beginning and end of the group curve is based on significantly less data points than the middle range where all of the children overlap. The curves for articles, modals and pronouns look like stochastically deformed S-shaped curves.

If we compare the graphs with Wells’ curve of chapter 4, we see that the deformed graphs in Figure 5.4 are not as smooth as Wells’ curves. This is quite understandable: there are only 6 children in this study, in Wells’ study there are over a hundred children, making the curves smoother.

In order to test this assumption of S-shaped curves, the non-linear model for S-shaped growth (van Geert, 1991) was compared with the classical linear regression model. Table 5.10 shows the results of the model fits. The least squares of sums measures the distance between the data and the model. In all cases the growth model produces a considerably better fit than the linear models.

Least squares	articles	modals	pronouns
Growth	0.25	0.38	0.51
Linear	4.0	3.6	3.2

Table 5.10. Least squares of sums of fits of the growth and the linear model

The disadvantage of age averages becomes apparent if we look at the example in table 5.11. Suppose that seven children were measured on some developmental scale or variable. The individual changes of 7 children show a jumpwise change, from 0 to 1. Thus, the change from 0 to 1 is instantaneously, and there is no regression back to 0. If all children jump at a different age, which is conceivable for a number of developmental changes, e.g. children start to walk at different ages, the average-on-age shows a line. So, the individual patterns are lost if data are averaged over age (here: weeks, months or years). It is clear that if the jumps were aligned on age, and averaged subsequently, the jump would still be present.

It is interesting to take a look at the estimated growth parameters of the empirical data, especially the exponent p. When p = 1, the growth model produces the classical S-shaped curve. With increasing p, the resulting model goes after a slow start towards an increasingly steep, cliff-like curve (van Geert, 1994, 1995). Table 5.12 shows that the average value of p is almost 1. That is, the “average” curve of function word increase is an ideal S-shape (note, however, that there are considerable differences between the variables). An average over age has also been applied to the empirical data. The results of the average over age are in table 5.12.

Age	child 1	child 2	child 3	child 4	child 5	child 6	child 7	Average
1	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	1/7
3	1	1	0	0	0	0	0	2/7
4	1	1	1	0	0	0	0	3/7
5	1	1	1	1	0	0	0	4/7
6	1	1	1	1	1	0	0	5/7
7	1	1	1	1	1	1	0	6/7
8	1	1	1	1	1	1	1	1

Table 5.11. Average over age: a jump becomes a line.

	Articles	Modals	Pronouns	Average based on age
Initial level	0.0132	0.0001	0.00741	0.0069
Exponent	1.40258	0.76545	1.12049	1.09617
Growth rate	0.27496	0.06668	0.15021	0.16395
Carrying capacity	0.61102	0.49902	0.63663	0.58222

Table 5.12. Growth model parameters for the age averages

It may be concluded that the average-by-age data are best described by a classical, continuous S-shaped growth model. However, as explanatory models, growth models of language apply only to individual data (van Geert, 1991). Averaging across age may in fact conceal interesting dynamic properties of individual growth patterns. A set of differently timed sudden jumps, in the form of a steep “cliff”, for instance, produces a perfectly smooth S-shaped curve if the data are averaged across age (see table 5.11). It is clear that, in this particular example, the resulting smooth S-shape misses the essential feature of individual growth patterns, namely the existence of a steep “cliff”, the sudden jump. In the next section I shall discuss the individual growth curves aligned on the jump. Suffice it to say for now that the averages of the parameters based on parameter estimations carried out with the individual instead of the group data show a significantly different pattern. The rate and exponent - which determine the form of the S-shape - differ considerably.

Data aligned to the jump

The problem with age differences is that they eventually conceal individual differences in the timing or onset of development, for instance, in the onset of the jump towards the rule-governed use of function words. Thus, instead of taking age as reference point, the data from the children should be aligned and averaged on the basis of the process characteristics themselves. That is, I must first align the data such that the first major jump, if any occurs, in the data of all the children coincides, and then compute the resulting average scores. The hypothesis is that, if I do so, the averages will show a pattern of change consisting of an initial phase where the frequency of function words is close to zero, followed by a sudden jump, followed by a pattern of stable proficient use of function words.

The problem with calibrating the data on the basis of a common sudden jump is that

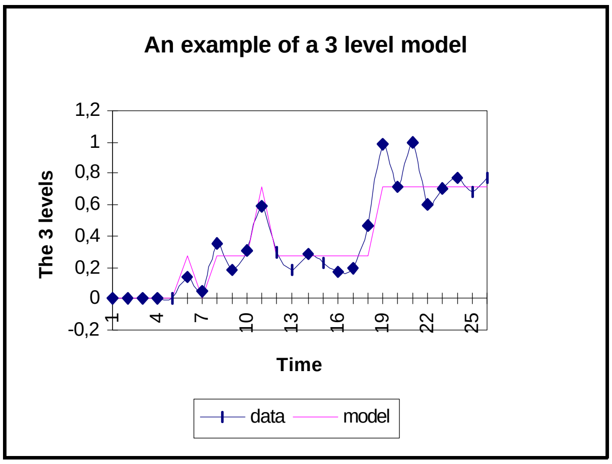
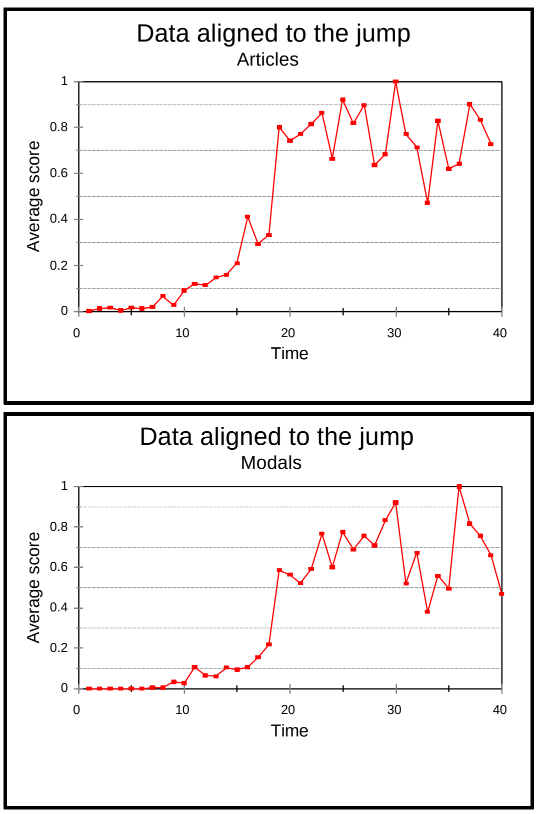


Figure 5.5 A three level model.

not all the data sets show a clear jump-wise increase. In order to overcome this difficulty, the individual data series were fitted to a three-levels model. Three levels were chosen on grounds of the multimodality analysis, according to which there are three equilibria in the data (although the evidence was not very strong). This three-levels model consists of an initial, an intermediary and a final (maximal) level. The minima, media and maxima were chosen so as to minimise the value of the sum of least squares of the differences between the empirical data and the three-level model. The model that was fitted to the data had the following form. Assuming that $(x_1, x_2, \ldots)$ is a data set (e.g. the number of articles counted during each observation of a particular child), I estimated three parameters Min, Med and Max, such that the sum of squared differences between x_t and y_t is minimal. The function for the estimation of parameter values had the following form:

$$y_t = \{ \quad \begin{array}{l} \text{Min if } |\,\text{Min} - x_t\,| < |\,\text{Med} - x_t\,| \\ \text{Med if } |\,\text{Min} - x_t\,| > |\,\text{Med} - x_t\,| \text{ and } |\,\text{Med} - x_t\,| < |\,\text{Max} - x_t\,| \; \} \\ \text{Max if neither Min nor Med condition applies} \end{array}$$

The estimation procedure was carried out with *Magestic*. Figure 5.5 shows such a three-level model on the basis of an example.



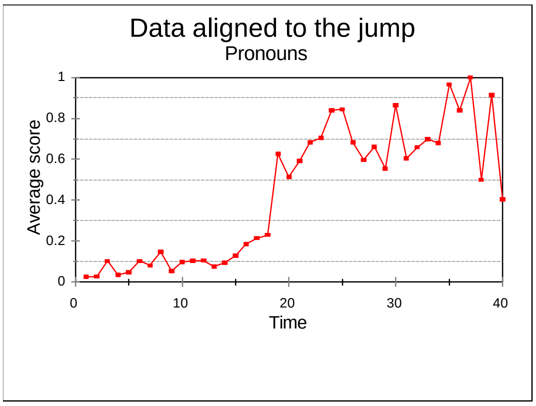


Figure 5.6. Average over jump.

After fitting the individual curves, I determined the first major jump as the point at which the curve touched its own maximum for the first time (the maximum being the best fitting estimated maximum value). All individual curves were aligned for each of the closed class categories (articles, modals, pronouns) such that the first major jumps fell on the same point. The averages of each point were computed by dividing the sum of the values at that point by the number of children at that point. Figure 5.6 shows the resulting curves, table 5.13 shows the average parameter values of the three variables averaged over the jump.

I tested the assumption that these curves represent sudden jumps (not discontinuities, but a rapid increase) by estimating their growth parameters. Table 5.12 gives an overview of the parameters. The exponents and growth rates indicate whether the curves are sudden jumps or not. If the exponent approaches 2 and / or the growth rate approaches 1, the resulting growth curve takes the form of a steep cliff-like change, i.e. a sudden jump (see van Geert 1994).

If the average parameter values are compared with those based on the individual curves (see Table 5.1), we see that the most important parameters, the exponent and the growth rate, are of a similar order of magnitude (exponents are 1.41 and 2.07, growth rates are 1.68 and 2.04). The aligned data, however, show the jump-like increase in a much clearer fashion. The most important parameter in this respect is the exponent: if it is equal to 2, the resulting growth model is the so-called cubic model (van Geert 1994) which produces a strongly accelerated growth in the form of a steep, cliff-like increase.

Growth model fit	articles	modals	pronouns	average
Initial state	0.02	0.02	0.06	0.03
Exponent	1.69	2.06	2.46	2.07
Growth rate	1.31	3.00	1.80	2.04
Carrying capacity	0.6	0.60	0.71	0.64
Least squares				
Growth model	0.45	0.24	0.36	0.35
Linear model	4	3.8	3.2	3.67

Table 5.13. Estimated growth parameters for the data aligned on the basis of first major jump and a comparison between the growth and the linear model by means of sums of least squares

The growth parameters for the data sets of the six children aligned on the basis of the estimated first major jump clearly show that the pattern of increase in the frequency of closed class words is similar to that of a cubic growth model, which is characterised by a strong and sudden acceleration of the growth rate. The least squares show that the growth model is 10 times as good as the linear model, i.e. the deviation is as 10 times as high for the linear model.

There is, apart from the linguistic connection, i.e. function words refer to functional categories, a developmental connection between the three variables. Function words start to grow in the same age span. These development of these words is individually timed, they are not generally age related. According to the fits of the individual growth curves and the averages over the jump, they grow suddenly. This leads to the conclusion that the assumption that language development is best explained by age and by a linear model is at least dubious.

Summary of the continuous analyses

Averages, both over age and over jump, yield evidence for the fact that in the case of these averages there is no linear increase. Age averages yield s-shaped graphs, jump averages yield cubic graphs. The parameter estimations based on the six individual children allow us to show the difference between a growth model based on averages of those individually estimated parameters and those of a growth model whose parameters were estimated on the basis of the age averages. The difference is striking: whereas the average-by-age parameters show a slow increase, the average-by-individual parameters produce a steep, cliff-like curve (Ruhland & van Geert 1997).

I presented two sorts of averages. The *age* averages of the three variables, that resemble the method of Wells (1986), show growth curves that are s-shaped. In addition to these s-shaped curves, these averages show no individual trajectories with the dynamic characteristic that were found in the growth curves. This means that if one is to draw conclusions from age averages, one must be extremely careful.

Although the *jump* averages do not show the individual characteristics, they do show that if we let go of age as a prediction for development, and take another denominator, e.g. the jump in the data, the resulting growth curves show quite a difference with the age averages. The jump found in the individual data does not disappear in these averages, and therefore, this average is a better representation of individual development than age averages.

The three models that are applied to the averaged data (the jump model is only used for finding the levels and the first major jump in the data) show that an increasing complexity of a model also leads to a better fit. Interesting to note, however, is that the logistic model shows complex behaviour (regression, jumps, fluctuations, etc.) despite the fact that there are only 4 parameters. The logistic model is capable of modelling the dynamics of processes in development. Sudden increases, fluctuations, regression are inherent to this model, just as they are inherent to development.

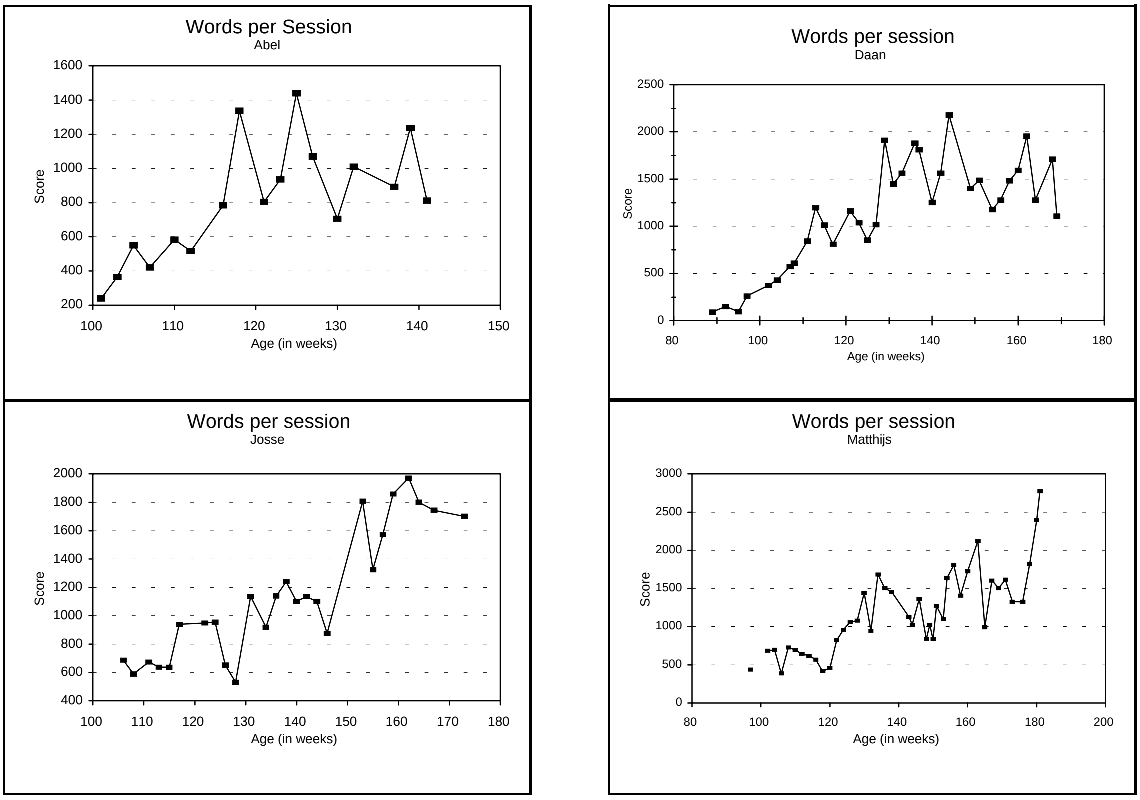
In sum, there is no evidence that the development of function words is best explained from a linear or discontinuous point of view. The non-linear fitting has implications for the study of language development. It does not mean that age averages are wrong. However, one must be careful with respect to certain claims about development. Group data do not yield a solid description of (individual) development. Only age dependent development (i.e. strictly timed development) will benefit from such averages. The second point to be made is that whatever the averages, in all cases the data are best modelled by a non-linear fit.

5.3 Explanations of the results found

Introduction

Although there is some evidence from the catastrophe flags that language development (i.e. function words development) is discontinuous, the evidence is poor. Whether the change is continuous or not, the least one can say is that it is a rapid change. This findings calls for an explanation: is the sudden change in the production of function words a selforganising process, or is it an epiphenomenon of a change in the child's productivity variables, or a change in the language of the parents (i.e. some form of imitation)? There are three different sorts of explanation. First, the production of words during a session might cause a change in the dependent variable. Put differently, a sudden change in the frequency of articles might be caused by an increase in the overall production of words (and utterances) in the speech of a child. Second, the length of sentences may increase as a consequence of other variables (see also chapter 6 for a discussion on these variables). This increase provides more syntactic slots, and as a result functional slots may come available. Third, a variable outside the child changes. For instance, changes in the production of function words in the parents might induce the sudden increase of function words in the children.

Dependency on productivity



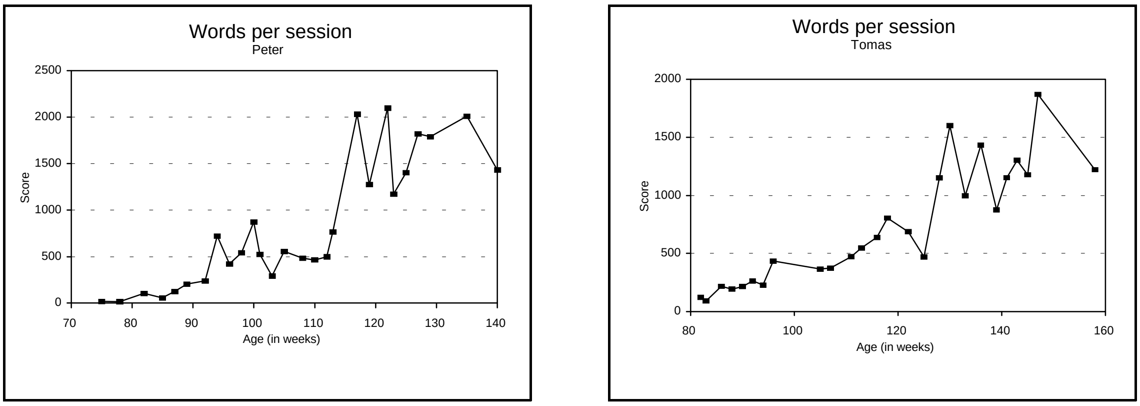


Figure 5.7 Words per session.

The number of words spoken during a session increases in an s-shaped fashion. That is, it is assumed that, on average, all word categories and word sorts will show that s-shape (see Ruhland, Wijnen & van Geert, 1995 for an s-shaped increase of finite verbs). This is important since one of the factors which causes a variable (e.g. function words) to increase suddenly could be the amount of "words per session", i.e. the language productivity. That is, if function words show an increase in the speech of a child, then this increase might be caused by an increase of the amount of words a child produces during each session, since function words are a subclass of words.

The graphs with the productivity per session shows individual differences. The children show stages in their language productivity, i.e. the amount of words per session remains temporarily relatively stable, i.e. with fluctuations, and after each stage, a child shows a considerable increase of words per session. Peter's language (i.e. words per session), for example, is characterised by three stages where the production of words in a session fluctuates around a equilibrium.

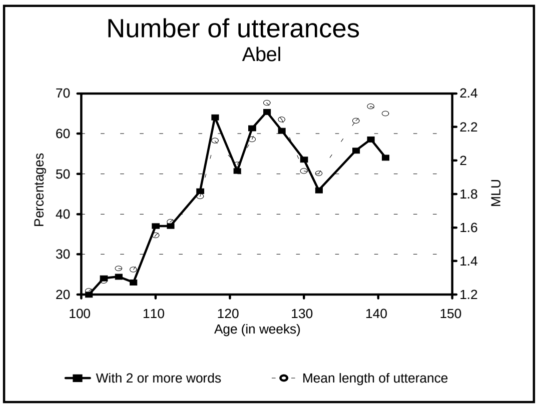
The question is whether the number of function words of each child increases due to an increase of productivity, or that it is a change that is independent of productivity. In fact, this is a classic 'The chicken and the egg problem'. I think that it is not possible to prove either case. It is not possible to determine the nature of the causal relationship between both the increase of the dependent variable and language productivity. The empirical data show an increase in the overall productivity and an increase in function words. It is assumed that the increase of the dependent variable (function words) helps to increase the productivity, and not the other way around, since the number of function words is very small (close to 0) when they start to grow, while the overall produtivity in all children has already reached a certain level (between approximately 500 and 1000 words per session). Although the increase in productivity is caused by an increase in function words **and** by an increase of non-function words (i.e. content words), it cannot be claimed that function words grow due to a growth of the overall productivity, since function words are a subclass of all words.

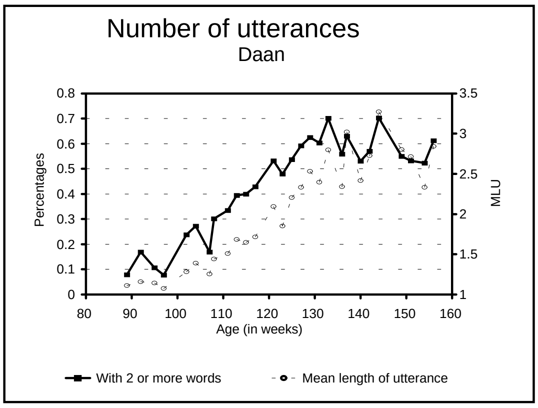
The crux of this section is that both function words and the total amount of words in a session show a coherence, but it cannot be decided whther either one explains the development of the other.

Length of utterance: means and proportions

The ratio of one word and multiword sentences is important since the amount of multiword utterances is an index of possible slots for syntactic structures (Ruhland & van Geert, 1998a; see also chapter 6). In figure 5.8, the ratio between these two kinds of utterances is expressed. This ratio is needed to decide when a child is lengthening his utterances, i.e. when a child leaves the one word stage and utters sentences that contain 2 or more words. I also added MLU. Although it is not a entirely satisfactory index (see for criticism Chabon, Kent-Udolf & Egolf, 1982), it is claimed that MLU should be over 1.75 before function words will appear (Brown 1973).

A few conclusions may be drawn from figure 5.8. First, in all children MLU is over 1.75 when function words start to grow. Brown’s obsvervation for English is also correct for the Dutch children in this study. Second, the graphs of multiword sentences and of MLU share a resemblance: the pattern of change is similar. The change from one word sentences to multiword sentences appears to be a quick change, but only three children (Matthijs, Peter and Tomas) have been recorded early enough to show such a change.





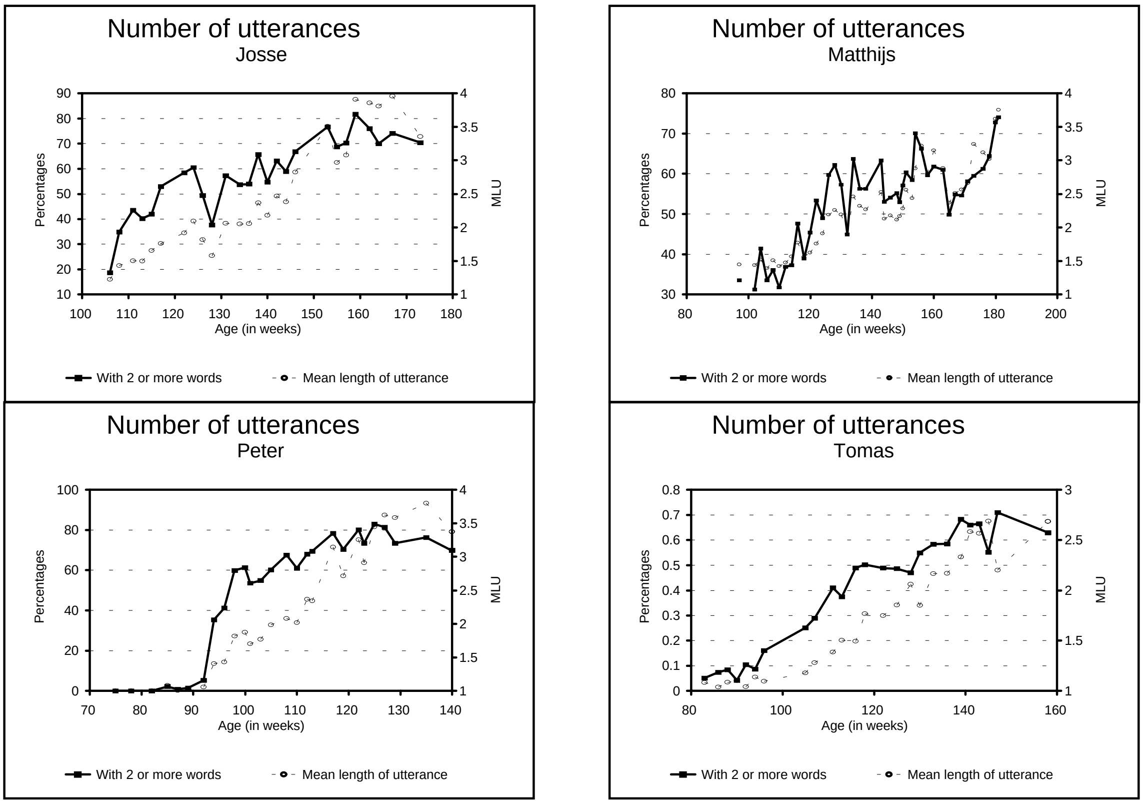
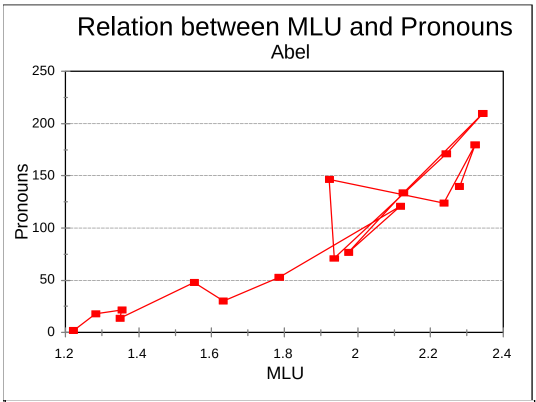
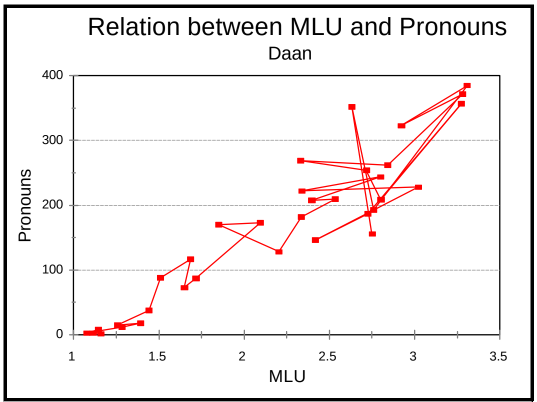


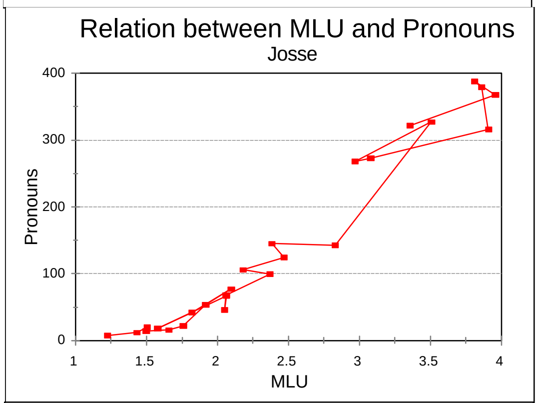
Figure 5.8. The length of the utterances: means and proportions.

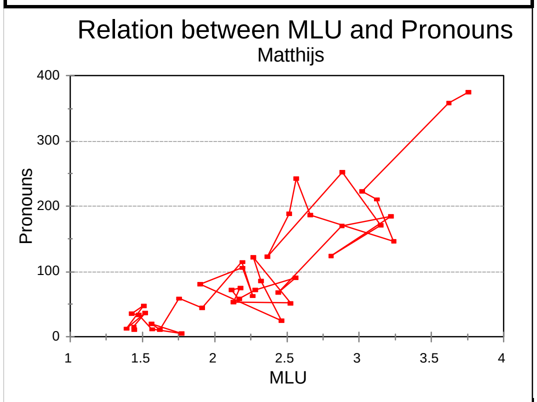
The multiword utterances are an indication that children lengthen their sentences. Whatever the reason for this lengthening is, it has consequences for the development of syntax. The reasoning is the following. Longer sentences, for example with two instead of one noun, provide extra slots in which syntactic elements fit naturally. This means that if a child increases the length of an utterance for any reason (see chapter 6 for a discussion on causes of these changes), it most likely will lead to the introduction of functional categories. Furthermore, MLU looks in all children very similar to the ratio of multiword / one word sentences, and children develop function words when MLU is in fact over 1.75.

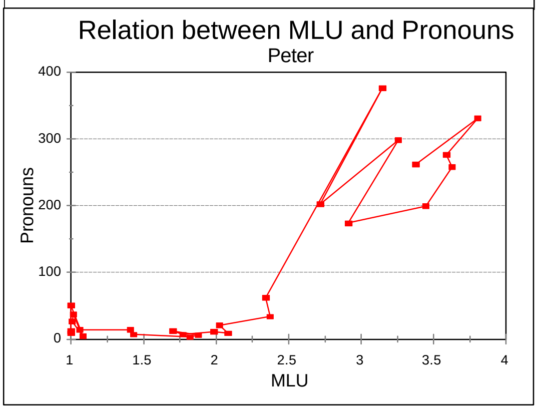
I argue that the increase of sentence length is not the result of the introduction of function words or functional categories. It is the other way around: longer sentences are attractors for functional categories. This can be concluded from the following reasoning.

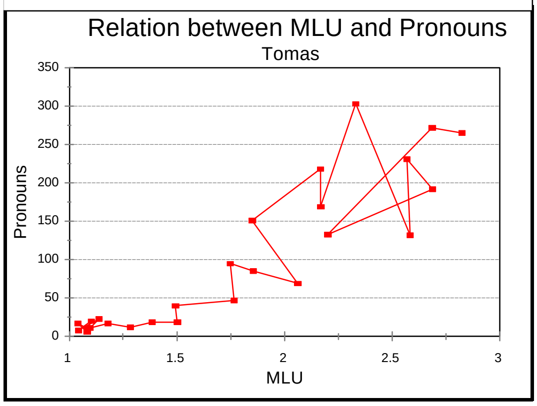












The introduction of function words is relatively late. By the time of that introduction, most utterances contain on average 2 or more words. These words must be content Figure words, since there are no function words yet. Although function words will increase the sentence length, the increase of length is a consequence of more content words, thus lexical items. These lexical items call for more syntactic slots, with as a result an increase in functional categories (since most lexical categories already developed), and therefore, an increase of function words. These words **follow** the lengthening of utterances, or in other words, longer sentences are an attractor for function words. This is expressed in the following graphs. They contain a state space of

MLU and function words. I choose the pronouns again to elicit the coherence between the length of the utterance and function words.

What these graphs illustrate is that when MLU increases, the number of pronouns increases. However, this increase is (curvi-)linear. Although this may be a demonstration of the importance of an increased sentence length as a condition for function words to appear, these graphs are no explanation. That is, MLU cannot explain the development of pronouns.

The data of the parents

The data of the parents show a certain baseline (e.g. the amount of functin words during a session) that adults have in their normal speech to children during one hour of observation. This baseline is almost reached by the children by the time they use function words regularly, i.e. after the sudden increase.

The most important aspect of the parent’s language is, however, not their actual level of production, but their change over time. This is the reason that these data are not presented in graphs here. If the parent’s language is responsible for the increase in function words, a jump in the child’s use of function words must be preceded by (or at least coincide with) the jump in the parental frequency. The “parental” factor can be checked by using the frequency scores of the parent(s), that are based on the corresponding sessions of the children. First, I assume that a child does not change his language in concordance with the language of the mother. There is no direct relationship between the data of the mother and the data of a child *within* one session. Instead, the time lag hypothesis is adopted: the child’s language resembles the mother’s language with a time lag. Thus, the analyses refer to a time lag, i.e. the language of a child follows the language of the mother of a few weeks earlier. Since it seems not plausible that the child acquires a language from one specific moment in time, but more likely, to a period of time (several weeks), the analyses are based on an average of the last 3 recordings. This number may be arbitrary, but it turned out that an increase in the amount of weeks of the average does not change the average. Each average is very close to the individual session scores of the mother, i.e. the mother scores per session do not vary greatly over time.

The purpose is to find out to what extent an age effect may be regarded as the main cause of change, and to what extent the mother does or does not contribute to the language of the child. In other words: the analyses must yield evidence for or against the fact whether the language of the mother is the main reason for the sudden change in the development of function words of a child.

The procedure in testing this hypothesis is that I fitted model on the data of the children. The model that is used for the multiple regression analyses is a polynomial curvi-linear model. With this model I fitted the child's time series (i.e. the age effect), and subsequently, the influence of the mother (i.e. the input) is fitted. So, the first step is to determine the age effect of the child, and next the influence of the mother's language after the age effect has been 'removed'. In the following tables I summed up the p-values of the multiple regression analysis. Note that I used a linear model, not a logistic model. The reason is that on the basis of the hypothesis that the parental language influences child language supposedly in a linear fashion, i.e. during the phase of rapid growth. I will show that a multiple regression analyses explains the development of function words as an input independent process.

Model	Abel	Daan	Josse	Matthijs	Peter	Tomas
Linear	0.07	0.00	0.23	0.00	0.00	0.23
Quadratic	0.07	0.00	0.41	0.00	0.00	0.1
Mother	0.78	0.22	0.46	0.27	0.57	0.88

Table 5.14. The relation between the child's and parent's function words: articles.

Model	Abel	Daan	Josse	Matthijs	Peter	Tomas
Linear	0.00	0.00	0.58	0.98	0.1	0.79
Quadratic	0.00	0.00	0.31	0.61	0.1	0.47
Mother	0.26	0.00	0.00	0.07	0.76	0.37

Table 5.15. The relation between the child's and parent's function words: modals.

Model	Abel	Daan	Josse	Matthijs	Peter	Tomas
Linear	0.99	0.031	0.00	0.05	0.35	0.79
Quadratic	0.89	0.19	0.00	0.02	0.24	0.43
Mother	0.19	0.18	0.54	0.89	0.38	0.21

Table 5.16. The relation between the child's and parent's function words: pronouns.

Each table has to be read from above. Thus, the tables show to what extent the llinear model explains the relationship, how much more the quadratic (linear) model adds to the analysis, and to what degree the mother makes a significant contribution to the change in language of the children. In two cases the mother contributes to the development of the function words, i.e. the modals of Daan and Josse. In all other cases, the data of the parents show that a parental factor is probably not the reason for the sudden change in the language of the children. There is a constant input: the linguistic productivity in terms of function words of the parents does not increase or decrease during the period that the child suddenly increases the production of function words.

Apparently the development of function words is a process that does not simply follow the parental input. Although the input is very important for the development of a language in a child, the conclusion must be that the child is relatively independent in his quantitative development of language. In the next chapter, I argue that the explanation and the interpretation of the data can be based on selforganisation being the result of linguistic development itself.

Summary

I offered three possible explanations for the change found, based on a model of independent variables (i.e. why the change in function words is sudden). First, the function words increase due to an increase in other variables (i.e. the change is caused by other independent factors). The best candidate presumably is the amount of words spoken during a session. Analyses, however, suggest that the words per session do increase over time (the production of speech increases as the child becomes older), but there is no coherence between this independent variable and the sudden change in function words. Thus, the increase of words per session is probably not the cause of the increase of function words. It might be the other way around: the productivity increases due to an increase in function words and other variables. The problem is a ‘chicken and egg’ problem. It is likely that much of the increase of the productivity is caused by the introduction of function words.

Second, the increase of utterance length would call for functional categories. That is, the lengthening of the utterances (expressed by an increase in both MLU and multiword utterances) is the result of an increase of lexical categories. Since lexical categories like nouns and verbs call for functional categories, longer sentences make syntactic slots become available. Longer sentences and the availability of syntactic slots can explain the timing of appearance of function words. However, they cannot explain the suddenness of growth of function words.

Third, the parents might increase their use of function words, and children follow the parental increase. This would indicate that a child uses the input to imitate it directly or with a time lag. However, on the grounds of the analyses, this assumption must be rejected. The multiple regression analyses do not confirm this hypothesis. The increase of function words in the language of a child is solely the result of an age effect. Thus, timing nor shape of change are the result of a change in language of the parents.

6

The changing tide: the discussion

And the power of the changing tide
blinded by the night
The sweet sound of the changing tide of your nature

Your nature by **Hothouse Flowers** from their CD **An emotional time**

6.0 Introduction

Nearly 40 years of study of child language using modern linguistics has brought up one of the richest areas in linguistic research. However, I argued that little attention has been given to *processes* in language development, i.e. quantifications over an index of time (chapter 2). I introduced *non-linear theories* that either aim at explaining growth (continuous change) or that offer indications for discontinuities (chapter 3) and a method for the study of language development using *dense time series of longitudinal data* (chapter 4). The *results* of the analyses were presented in chapter 5.

The interpretation of the results is directed to the main research question I presented in chapter 1, and which was elaborated in chapter 4. The question whether language development is continuous or discontinuous, is discussed in the light of evidence from this study, i.e. from the analyses and the interpretation of the data. I link this evidence to both linguistic and psychological theories on language development (cf. Ruhland & van Geert, 1998a, Ruhland, Wijnen & van Geert, 1995, Cats & Ruhland, in press, Ruhland, 1998a), and with other empirical findings from the literature.

The method in this study is one of the ways to study quantitative language development, and therefore, some questions cannot be answered. For example, longitudinal spontaneous language does not allow one to test the remaining flags from catastrophe theory, especially those that call for a perturbation. I discuss improvements to a further study of non-linear development, and I plead for a study with an interdisciplinary character.

Finally, I discuss the relationship between DSM and CT, and a third non-linear theory and a collection of models, Connectionism. The goal is to outline a model for language development that is not build on innate structures or rules that prescribe development.

6.1 Flags and fits: discontinuous or continuous development?

At the end of chapter 4, I stated the following research question:

> Under the assumption that, according to the literature, function words are introduced very rapidly (at least) and that this reflects the succession of two stages (a telegraphic stage and a differentiation stage), the question is if there is any evidence of a catastrophic change (defined by the flags) in the development of function words in the six children between 1;6 and 3;0 years of age, measured on the basis of frequency analyses in a longitudinal design using two weekly observations.

The discussion in this chapter is preceded by a recapitulation of chapters 2 and 3. The theoretical starting points are discussed in the light of the arguments in favour of either discontinuity or continuity. This will lead to a discussion of the data in relationship with existing linguistic and psychological theories on the empirical and the theoretical level. Explanations of the results come in terms of the shape and of the timing of function words in child language.

Theoretical observations

Theories on development offer a limited prediction on the shape of development. Human behaviour is not exclusively determined by the environment (e.g. the parents), neither by innateness (e.g. genes). With respect to *change*, I distinguished two sorts of theories. *Non-Stage* theories predict linear or at best gradual change. *Stage* theories offer the possibility of sudden change, although sudden changes may be disguised by other factors. Furthermore, I distinguished continuous (linear and non-linear) and discontinuous change. If development is continuous at bottom, change cannot imply stages. However, the definitions of discontinuity and stages are (at least) fuzzy. Glaserfield and Kelly's (1982) ideas were used to define a stage, and I reformulated their ideas to the definition of 'a change on another dimension than time or age'. However, this definition is not sufficient. It is unclear what the boundaries of a stage are, and the stage definition lacks a structural account. Transitions from one to the next state are mostly sudden, since a gradual change implies that stages cannot clearly distinguished, unless there are a sufficient static properties in stages. A structural theory has to decide how these static properties differ from real stages.

With respect to *language*, I discussed theories on language and language development. Linguistic theories offer explanations for the variety within a language, and for the links among languages. Developmental predictions of linguistic theories are limited, but linguistic theories do have the possibility of explaining structures in language. Underlying structures, traces and movement rules explain the syntax of a language. Features in the syntax of many languages including Dutch are functional categories. They are abstract, syntactic, and important elements in a sentence.

One assumption in continuous linguistic theories is that children have a body of linguistic knowledge called UG. The question is whether UG principles are available to children from the beginning or not, i.e. whether the development of syntax is, in the linguistic sense of these words continuous (all principles of UG are available) or discontinuous (some or all are not available). Generative grammar theories assume that development is guided by principles that explain the order of change. A precise quantitative analysis on the basis of UG is not possible. Change is either a gradual process, e.g. an extremely rapid development (in which the continuity hypothesis is obeyed), or a discontinuous one (UG principles are not present from the beginning). UG makes it plausible that development is constrained by general linguistic rules, that children and adults may learn other languages, and that the learnability problem is solved by postulating innate syntactical knowledge. It is assumed that this knowledge is not specified as to the details of a language, but that it increasingly becomes specific when a child grows up.

A different approach to the study of quantitative paths, continuous or discontinuous (i.e. the psychological meaning), is found in non-linear models. *Dynamic systems* (continuous) models have been proposed by Thelen and Ulrich (1991) and van Geert (1991) for describing processes in motor and cognitive development. Instead of linear models, their models are based upon non-linear ones. Following van Geert (1991, 1994a), I assume that development is a process of dynamic interactions between growers. The dynamics are also expressed by a sensitivity to small changes in the initial values that can eventually lead to big differences in the outcomes. The language system that develops sets out constraints. These constraints lead to restrictions on possible forms of change of a psychological domain. There are two general applications of dynamic systems models. First, I use the models for fitting growth curves. Second, non-linear equations are used to model language development. In this chapter, I will add a linguistic analysis to my data and van Geert's analyses and models. *Catastrophe theory* (i.e. a discontinuous theory) offers criteria that may be employed in order to find discontinuous change. There are eight so called catastrophe flags, associated with specific kinds of empirical data. These flags have to be found in order to demonstrate phase transitions.

Empirical findings

Two opposite approaches to development have been used to answer the following question.

What are the paths of language development (in quantitative terms)?

Discontinuous models are applied in two ways: the use of catastrophe flags (the catastrophe detection method) to get indications for discontinuous behaviour, and a direct fit of a cusp model (*Cuspfit*; Hartelman, 1996), which is based on Cobb (1981), to test three models of development.

Evidence for discontinuous change comes from the catastrophe flags. Out of 8 flags from catastrophe theory, there is evidence for three flags. I found a *sudden jump* in function words. This flag indicates that a variable jumps from one equilibrium to the second. There is also *multimodality*. This flag shows that there is more than one equilibrium (the presence and non-presence of function words). *Anomalous variance* indicates that if there is a transition (or discontinuous change), the variance increases temporarily in an anomalous way, but this flag was found in only two of the six children.

Supplementary evidence for discontinuous change comes from a method of fitting several models to the data by using Cuspfit. Taking (cross-sectional) data as an input, it tests three models (linear, logistic and cusp). The results show that in all cases (articles, modals and pronouns), the cusp model provides the best model. A comparison of three models (linear, logistic, and cusp) shows that the cusp model is the best model to fit the data (according to cuspfit).

Fits based on growth models (van Geert, 1991, 1994a) show that equations with high growth rates (> 1) and a high power (> 1) yield the best fit. These equations are closely connected to catastrophe theory by the power of the equation, and the resulting curves show a rapid and cliff-like increase. That is, if the power of the equation comes close to 2, and the growth rate is high (e.g. close to 2) the resulting growth curve takes the form of a sudden, cliff like form, accompanied by fluctuations, and a low initial state.

The conclusion must be that, from the discontinuous analyses, it appears that the quantitative development of function words is rapid. There is some evidence for three catastrophe flags, the analyses with *Cuspfit* suggests additional evidence for a catastrophic change (i.e. the cusp model), and there is more than one mode (or equilibrium) in the data of children according to the analyses with a *kernel density technique* and a *mixtures distributions analysis*. There exists large individual variability in the development of function words in each child. The literature suggests that function words develop suddenly. Using dense time series I showed that there is indeed a sudden jump in at least some of the children. However, there is insufficient evidence to support the hypothesis of discontinuous change and simultaneously reject the continuous change model.

The next step was to test the continuity hypothesis. *Continuous* models test development with respect to the general linear or even non-linear (but not "discontinuous") trend. Continuous models may be linear, polynomial, or logistic with on the x-axis the independent variable (e.g. time or age), and on the y-axis the score of a variable. The important characteristic, however, is continuity in a mathematical (i.e. psychological or quantitative) sense. This does not imply that the psychological model needs to be continuous: a series of discrete learning events could lead to continuity on the mathematical level. For every point in time, there is a score on the y-axis. There are several tests for testing the continuity hypothesis. I tested two models: a linear and a growth one. The conclusion must be that a fit with growth models to dense time series of individual development is much better than a linear fit.

Finally, I presented two sorts of averages to show that age averages hide the dynamics of development. Averages over age show similar patterns as those of Wells (1985), namely s-shaped forms. However, individual paths show that there sudden changes in some of the children. Dense *individual* time series with two-weekly recordings show that there is at least a sudden change in the use of function words. Averages over the jump, the biggest difference in a set of data, shows, under the assumption of a variable tied process, not an age tied process, that the jump found is still present, which is an extra indication of sudden change.

I offered explanations for the change found, i.e. why the change in function words is sudden. First, the function words might increase due to an increase in other variables, i.e. other independent factors. The best candidate presumably is the amount of words spoken during a session, i.e. language productivity. The words per session increase over time: the production of language increases as the child becomes older. However, it is not possible to determine the direction of the causality. The only to be made is that there is a interrelationship between both language variables. Second, the length of sentences increases in the same age span that function words arise. This fact correlates with the assumption that longer sentences provide slots for function words and functional categories. Third, the parents might increase their use of function words. This would indicate that a child imitates the input directly, or with a time lag. However, analyses do not confirm this hypotheses. The increase in function words is not directly related to the language of the parents, i.e. the input.

6.2 Quantitative and qualitative explanations of rapid change

We have seen that the change found is not caused by the parents or by an increase in the child's speech productivity. The question that I shall discuss in the present section, is why, if no such external factors suffice, function words show the jumpwise increase as found in the data. I discuss psychological and linguistic theories with respect to behavioural versus underlying variables in the light of the sudden change. The explanations come from two directions. First, I present a quantitative discussion on the shape of change. Second, I discuss the relatively late appearance of function words.

The shape of change: quantitative explanations

A dynamical systems explanation

A dynamical systems approach enables us to clarify why a sudden jump occurs. In a growth model of language variables such as articles, modals and pronouns are various syntactic structures that are considered connected growers. That is, they follow a path of mutually related development. I mentioned earlier that the growth in the use of certain syntactic structures supports the use of pronouns, because these structures provide 'slots' into which pronouns fit naturally. On the other hand, growth in the use of pronouns supports the search for syntactic structures in which the placement of pronouns is syntactically governed. Thus, one reason for jumps might be that a variable (e.g. closed class words) gets support from another grower, namely the development of longer sentences with inflected verbs, which results in closed class words growing faster.

In addition to these relationships of mutual support, the disappearance of a competing relationship (e.g. competition with two-word phrases) may lead to a sudden increase in growth rate and a new equilibrium. Thus, a sudden jump may arise because the competition with another grower has disappeared. The grower that jumps to another level of equilibrium no longer suffers from the limitations set by another grower. It may now use all the available resources eventually leading to a jump.

A third explanation is that a precursor relationship, for example, the inflection of verbs, causes a new grower to emerge. In terms of development, this precursor releases certain structural possibilities of the language system. In our model (Cats & Ruhland in press), this relationship is implemented as the coproducer of new syntactic structures. Precursor relationships are 'pushers' of new variables. Since the precursor that was chosen on theoretical grounds sets free functional categories, a very rapid development is possible, because these functional categories, unlike the verb inflection rule, do not call for rule application across lexical items (as in the case of verb inflection) in the case of function words. These categories are filled in by function words in a relatively simple way.

One of the assumptions is that growers in a system or model are forced to grow because there are *attractors*. Attractors are, in terms of thermodynamic theories, points in the developmental landscape that possess less energy. This means that a variable or grower is attracted to the low energy state. This state is not prescribed in a system, i.e. in the language itself. It is a result of the interaction between words, i.e. the process of a grammar. This can be called self-organisation: the organisation of grammatical features is inherent in the way linguistic theory argues that relationships between syntactic elements are organised. This organisation between syntactic features must be acquired in development, and during that process, attractors are states in development. For example, the first attractor in development is vocabulary, i.e. the knowledge of types of words. These words children acquire are lexical items like nouns and non-inflected verbs that refer to entities in the child's environment. The attractor for function words is at least the length of the utterance. This attractor implies that if an utterance is long enough, i.e. a multiword utterance (e.g. with MLU over 1.75 words per sentence), function words appear. However, I think that this is not viable explanation. There is from a syntactic point of view no strict reason that function words should appear as soon as sentences are long enough. Below, I argue that finite verbs are a more attractive candidate to serve as an precursor for function words than the length of sentences.

The shape of change does not only need to explained from a dynamic systems point of view, it also needs to pay attention to individual differences.

Pronominal and nominal styles: individual differences

"Don't you know that it's different for girls?" Joe Jackson, **I'm the man**.

In chapter 4, I raised the question whether the distinction between pronominal and nominal (Bloom, Lightbown & Hood, 1975) or referential and expressive styles (Nelson, 1975) enables us to make predictions on the quantitative change we might expect. I argued that there is no compelling reason to believe that a linear, a slow gradual or a very rapid change follows logically from either of the styles that according to Nelson (1975) and Bloom et al. (1975) exist in language development (Ruhland & van Geert, 1998a). Pronominal children may acquire pronouns early, but this development may be either very rapid, slow, or linear. In a similar reasoning, nominal children are late, but when they catch up they may show any quantitative behaviour (from linear to sudden change). The style of development, so to speak, does not rule out any of the kinds of change proposed. The answer to the question whether the pronominal-nominal distinction matters is that there is only a difference in timing: pronominal children are earlier (in age) in their acquisition of function words, nominal children will catch up later.

Children acquire their basic grammar and language in a few years. But although children normally acquire all aspects of syntax, they are not obliged to use function words like pronouns in the early stages of development. It means that children eventually develop different individual styles (Richards, 1990, Shore 1995, Bates, Bretherton & Snyder 1988). If humans are not obliged to use deictic or referential structures in communication, then every individual development is characterised by a different style of language, although they all end up with the same grammar. Do styles influence quantified development, i.e. the shape of change? There is, like the predictions of the linguistic theories in chapter 4, no reason to believe that styles are directly related to developmental shapes. Presumably there is only an influence on the timing of function words: pronominal children are earlier (in age) in their acquisition of function words, nominal children will catch up later.

Therefore, the data do not allow for an interpretation of the growth curves in terms of styles. The “choice” of either a nominal or a pronominal style by each individual does not automatically lead to the conclusion that a catastrophic change is impossible. Children may follow two possible paths. A child like Abel may differ from other children, but the linear pattern of his development may be explained by the fact that the function words in Abel’s language have not reached the second equilibrium at the end of the recordings, and that there is no “function-words-free” start, i.e. a considerable number of function words can already be found during the first observation.

Another important matter are the subjects in the study. Due to a curious coincidence, all children in the study are *boys*. A dynamic systems model of development leaves open the possibility of individual paths. There are some claims that the timing of girls is different from boys. *Qualitative* differences between boys and girls have been described by Huttenlocher, Haight, Bryk, Seltzer and Lyons (1991) who found that the vocabulary spurt is earlier in girls. Most studies reveal that there are various qualitative differences between girls and boys (Maccoby, 1988). She describes research, in which differences between ‘boys in all-boy groups’, compared with ‘girls in all-girl groups’, have been found. For example, boys more often interrupt one another and more often use commands, whereas girls often express agreement with what another speaker has just said and pause to give another girl a chance to speak. Apparently, speech serves more egoistic functions among boys and more socially binding functions among girls. There are communicative differences, i.e. in speech acts (Serbin, Sprafkin, Elman and Doyle

1984). According to Fenson, Dale, Reznick and Bates (1994), differences between girls and boys are small, i.e. girls are ahead of boys. In sum, there are no qualitative differences between girls and boys expected with respect to the development of syntax, and quantitative differences exist only in the timing of the start of the growth of language. With respect to gender differences, age ranges between girls and boys may differ, but that this difference in timing is small. With respect to the shape of development, there is no research that suggests that there are gender differences in *quantitative* development, i.e. in the growth curves.

A threesome of data, non-linear theories and linguistic theories

Any data set, regardless the complexity of the set, the number and magnitude of fluctuations, etc. can be fitted with a simple or with a highly complex equation. The problem, however, is that fitting alone is an empty exercise. What is needed is a justification for both the quantitative and qualitative patterns of language development. The problem, in short, is: what is the relationship between the data, non-linear theories and linguistic theories?

In Ruhland, Wijnen and van Geert (1995), we discern, at a general level, three types of relationships between a structural theory of development and a growth model. The first possibility is one of *independence*. Apart from the assumption that growth “must follow a structural possibility of the cognitive system” (i.e. a grower must be present; van Geert, 1991: 4), the structural model and the growth model may be completely autonomous. The growth model is capable of describing the observed changes in performance without recourse to hypotheses on the quality of the developmental process that may follow from a structural model. In fact, this seems to imply that the growers do not entertain any relation to the notions of the structural model. Thus, the growth model neither supports, nor refutes the assumptions of the structural model.

Second, growth models and a linguistic theory are *compatible*. This comes in two varieties. First, it may be the case that the parameters of a growth model that was constructed inductively (i.e. with the single objective of obtaining a perfect fit with the data) can be meaningfully interpreted in terms of a structural model, and that this leads to (quantitative) predictions that concur with the predictions derived from the structural model. Alternatively, a growth model that was constructed *deductively*, i.e. by choosing its parameters on the basis of notions and considerations supplied by a structural model, is supported by the empirical data.

Third, growth and linguistic models may *contradict* each other. Again, we envision two versions of this situation. In the top-down scenario, a growth model that was informed by a structural model does not fit the data, whereas, possibly, a bottom-up (i.e. based on the data) model does. Alternatively, an inductively constructed model comprises parameters that can be interpreted as refutations of (some of) the core assumptions of the structural model. In either case, the bottom-up model (that was validated by its fit on the data only) may be thought of as the instantiation of a hitherto unidentified structural model.

Thus, there must be a relationship between (elements in) the equations and linguistic theory, because the construction of a dynamic model without any theoretical justification does not make any sense: it is nothing more than a mathematical restatement of several growth curves without any coherence. Structural change can be related to the equations in three ways. First, the carrying capacity, K, increases or decreases. This means that a system can grow to a different level. Maturation could be a cause of an increase of K. Second, supportive or competitive factors can push a system to a new level. In that case, the carrying capacity itself does not change, but other factors raise or lower the value of K. Third, competitive factors may avoid a grower to reach a new level, and consequently the system does not structurally change. In a structural model like a linguistic one, there is, with a competitive factor, basically no change on the underlying level. The syntactic (or more generally, linguistic) knowledge of some variable does not change due to competitive factors. The relationship between K and theories on linguistic development are probably that the theory sets out the constraints, and K follows these constraints set out by the theory. In other words, this is a ‘one way street’ from linguistic theory to K.

Suppose that a child learns the rule to use the passive. Consider the following example.

6.1 The standing stones of Callanish on the Isle of Lewis were erected five thousand years ago.

It is not necessary to learn the passive, since all sentences maybe put in the active form. With competition between two language structures, the stronger structure wins. If this is the structure that was already there, the underlying linguistic structure does not change. Thus, if the active form is the stronger structure, passives do not appear in a language, and only examples like in 6.2 appear.

6.2 Druids did not build Stonehenge.

One may question the existence of such competition in language (development). A discussion on this subject, however, would lead too far from the current subject of the section (see MacWhinney, 1997 for a discussion on the solution of the logical problem).

An important question is: how can we relate the structural developmental theories to quantitative notions? In Ruhland et al. (1995) we looked at a few of these linguistic developmental theories and possible predictions they cast on the sort of change. These theories are Maturation, Parameter Setting and Lexical Learning. One of our conclusions was that these structural theories and hypotheses fail to make exact predictions about the forms of change. However, they do provide hints about the particular directions in which change might take place. If, for example, parameter setting is the correct theory of development, we may expect sudden changes. I deliberately say ‘may’, because other factors like memory may hamper a sudden change. Lexical learning, on the other hand, predicts slow and gradual (or even linear) change. All these predictions can be related to a dynamical systems approach. For example, the carrying capacity (K), that is, the equilibrium level of a variable (where K might be the linguistic knowledge of a child), may change either gradually or suddenly. This K-change may occur due to an increase in, for example, memory or an increase of syntactic slots. The question is whether there is any justification in the data for any of these possibilities. Since I found a jump, lexical learning, with its hypothesized linear or gradual increase, appears to be a false linguistic hypothesis. All theories (parameter setting (Hyams, 1986), lexical learning (Clahsen, Eisenbeiss & Penke, 1994) and a range of maturational views (cf. Felix, 1992)) can be tested against the findings of a dynamical systems model. It is not clear yet whether these approaches (dynamical systems modelling and structural, linguistic theories) contradict each other, or whether they are in some way complementary. However, it is plausible that functional categories (e.g. function words) follow a path of mutually-related development, which is supported from a linguistic point of view by assuming that if functional categories are not available in development from the beginning (the reduced competence), these categories develop simultaneously on the basis of the Case module (see chapter 4).

With respect to the discontinuity and continuity debate in linguistics, the data are not suited for deciding for or against either one, since this debate calls for a different approach in which a qualitative reasoning is requested. The continuity-discontinuity debate is a theoretical battle that cannot be won with quantified data. In this study, it is assumed that the rules of UG are not violated, and that therefore linguistic continuity is preserved. What seems plausible is that the absence of function words at least suggests that functional categories are not visible in the language of children, and therefore, not present on an empirical level.

The timing of change: qualitative explanations

Explanations of a linguistic kind

To continue the linguistic reasoning of the previous paragraph, and bearing in mind that the rapid development of function words is not caused by independent variables like vocabulary, or the input (i.e. the parents), I will argue that a syntactic organisation could be responsible for the development found. This has to do with the order of change, but also with the shape of change, since the way new syntactic properties are introduced is decisive for the possible change and relationship between function words.

One of the assumptions is that early child language can be rewritten in the form of a linguistic tree, which is build of lexical categories (NP, VP, PrepP and AP). There are three approaches to early child language, that I introduced in chapter 4: functional projections are either completely absent (lexical-thematic hypothesis), partly present (reduced competence), or all available. For example, Clahsen et al. (1994) assume that there is one functional projection, namely F. The question is what this F could be. Let me start with an assumption, and then find evidence for it. The assumption is that Inflection, i.e. finiteness of verbs, is the first functional category. Finite verbs serve as a glue between other building blocks (mainly lexical items like nouns). Empirical clues have been provided by Vainikka (1993), Wijnen (1996), and Powers (1997). They provide some evidence that Vfin is the first F acquired. If the assumption proposed here, i.e. the first F is I(nflection), is correct, the functional phrase contains the inflected verb, and the basic tree looks like figure 6.1, where F is any functional category (IP), and the other abbreviations are lexical categories (see chapter 2). One or more of the categories in brackets, NP, PrepP and AP, may appear in early language, i.e. when only one functional category exists.

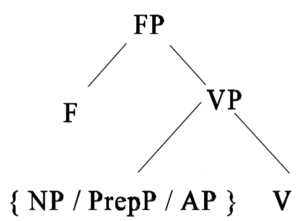


Figure 6.1. The syntactical tree before the introduction of function words.

With the introduction of functional categories (by means of, for example, function words), the tree ‘grows’. The growth of the tree means that branches are added. These branches represent new functional categories. Growth in terms of branch addition does not violate the linguistic continuity assumption. The introduction of function words and functional categories leads to a tree similar to that of Figure 6.2. The point here is the increased complexity of the child's syntax.

Box 6.1 Explanation of the abbreviations:

C = Complementizer, Agr = Agreement, S / SUBJ = Subject, T = Tense, O = Object, IO = Indirect Object, DO = Direct Object, P = Phrase, V = Verb.

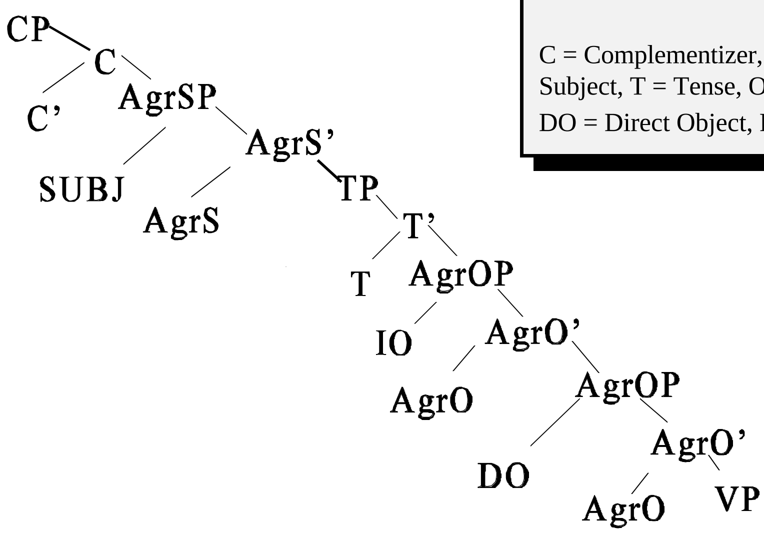


Figure 6.2. The syntactical tree with all functional categories (cf. Zwart 1993, 143)

Ferdinand (1996) shows that early child language does not lack all functional categories, but neither are they all present. Suppose that not all functional categories are available (see figure 6.1). New functional branches may be added all at once (e.g. as predicted by Parameter Setting) or one by one. Every new branch is instantaneously added to the tree. That is, there is a sudden jump in the equilibrium. Nevertheless, the effect in the behaviour may be gradual, because non-linguistic causes could be responsible: a slow increase of memory may hamper the instantaneous introduction of new linguistic elements.

The variables in this study belong to functional categories. Determinative pronouns and articles belong to the Determiner Phrase, personal pronouns take the position of the Specifier of the Complementizer Phrase, whereas modal verbs belong to the class of Auxiliaries Phrases. It is assumed that according to linguistic theory Case depends on inflected verbs. The introduction of finite verbs (Wijnen, 1996), i.e. IP, might release the class of other functional categories (see also Vainikka, 1993). This might also be true for language development, since in the case of Peter and Josse personal pronouns, for instance, does not start to grow until the verb inflection process is completed or close to completion. Other functional classes might be induced by similar syntactic processes. The conclusion is that finite verbs develop first, and this leads to the development of other functional categories, for example on the basis of Case distribution. This presumably is one of the mechanisms of change, inducing function words to grow.

On the basis of two empirical findings it is assumed that reduced competence is presumably the right hypothesis for early language development. First, finiteness (Wijnen, 1996) precedes the development of function words (Ruhland et al., 1995). In other words, these analyses confirm the reduced competence since finite verbs (a functional category) precede function words, i.e. there is no ‘all functional categories come in one go’. Empirically, there is support for the precursor relationship of verbs to other functional categories. Second, case distribution sets free a range of slots for functional categories. Theoretically, this means that the order of empirics is supported by a linguistic analysis.

In sum, the linguistic analysis serves as a formal model of qualitative change and as the account for the origin of linguistic knowledge. With respect to the quantitative findings of this study, however, neither linguistic hypotheses nor the ordering of the stages predicts or explains the shape of quantitative change.

A second issue concerns the question of the production as a function of comprehension. Criticism may hold that a child’s production of function words is the direct result of an increase in the comprehension of function words. Production may either follow comprehension, precede comprehension, or they may develop simultaneously. According to Petretic and Tweney (1977), comprehension precedes production. This implies that children do not have a problem of understanding function words. Other studies seem to confirm this finding. With regard to comprehension and production of word order in early child language, Roberts (1983) found that in Brown’s stage I (MLU 1.00 - 2.00) comprehension is verb specific, and comprehension precedes production. Studies on the comprehension of complex (in terms of syntax) sentences support the findings of comprehension preceding production, for example under the influence of phonological and rhythmic constraints (cf. Gerken & McIntosh, 1993), and although there are only few studies on the comprehension of function words by children, the conclusion may be reached that the rapid development of function words is not likely to be caused directly, for example, by a sudden change in comprehension, since comprehension is already completed by the time a child produces function words.

Transition mechanisms have had different interpretations and explanations. In Piagetian theory, for example, they were interpreted as change within the equilibrium model. In the case of conservation, for instance, a child can be forced into new conservation strategies because old habits do not function anymore. In the development of function words, there are no conflicts since function words do not fight for space (e.g. syntactic space or psychologically defined space like memory) with another variables. Nouns and pronouns (a nominal and a pronominal style) cohabit peacefully in a language. However, the lack of function words has been explained by the fact that function words have weak stress and that the omission of function words in telegraphic speech stage is caused by stress. Explanations of the early absence of functional categories and words often refer to some form of performance limitations. For instance, since function words receive only little stress, they are not easily perceived by the child and are therefore difficult to learn. Valian (1991), for example, claims that the omission of sentential subjects is best accounted for by performance factors. Gerken, Landau and Remez (1990) found something similar: the omission of function words is not caused by an encoding limit (i.e. problems with the input), but it is caused by phonological constraints. This evidence has been found in Dutch too (Wijnen, Krikhaar & den Os, 1994).

The phonological constraints might be the reason why children do not notice function words in the input of the parents. However, there is a problem with this assumption: why do children observe function words in a later stage (that is, in the differentiated stage)? The hypothesis is that an increase of memory leads to a bigger memory and attention span, which in turn leads to the introduction of function words. Empirical studies (e.g. Elman, 1993) are needed to turn the hypothesis into an answer, i.e. to relate the introduction of linguistic variables to non-linguistic issues (e.g. an increase of the memory span) when a linguistic analysis cannot explain that introduction.

Instability of the developing language system

There is still more evidence that the introduction of function words is not a gentle, gradual or linear process, but one that reactivates or even reorganises the system. The reorganisation is expressed in the difficulties of planning complex sentences. From the literature we know that the end of telegraphic speech is accompanied by non-fluencies. These act as indicators of instabilities of the changing system. The change of the system is expressed in an increase of productivity which is accompanied by disfluencies in the language of a child (Wijnen 1990). A child has problems in planning her utterances because there are new syntactical elements that make the utterances longer, but also more complex. The introduction of function words leads to planning problems which are not syntactic. This indicates that several ‘modules’ (from phonological to semantic, and also non-linguistic modules) are affected by the change in the system, which becomes less stable, or even unstable. Elbers and Wijnen (1992) found a relationship between the development of function words and non-fluency. Apparently, perturbances on one level (in terms of development of function words) leads to instabilities on the other (in terms of fluency). In other words, an increase of knowledge (i.e. the knowledge of how and when to use function words) is strongly related with the development of fluency. The disfluencies in the language of third year of life have been referred to as developmental stuttering: a temporary deterioration in the language of children.

Summary of linguistic explanations

The predictions and explanations of quantitative behaviour that follow from linguistic theories are very poor. Although these theories discuss the what of development pretty accurately, the how is a problem. This also means that a linguistic analysis can never be sufficient for the study of language development, under the required assumption that development is both a qualitative and a quantitative process.

Syntax trees provide a structural account for the production of linguistic structures. Since the system is not yet stable, which is indicated by the planning problems, this instability indicates a system in transition. None of the structural accounts explain the exact form of quantitative development.

Non-linguistic change

Neuronal development as a condition for sudden change

Another explanation of sudden development has been proposed in Wimmers (1996). Since all behaviour is based on neuronal activity, the development of language must relate to brain structures. One option is that the brain does not develop, but that is impossible. It is generally known that at birth cell formation and cell migration is completed, but that long-range connections are still to be established (Bates, Thal & Janowsky, 1992). Second, if brain structures develop linearly or gradually, then there is no one-to-one relationship between the sudden introduction of function words and the gradual brain development. Third, the brain may develop in spurts. In that case, language development might, as a result, also develop spurt wise. Evidence on the relation between neuronal structures, neuronal development and language development is scarce. Thatcher (1991) presents evidence that the brain develops in stages: neurological development is responsible for stages and discontinuous development. Although Thatcher has not related this stagewise brain development to age related language development, the first stage roughly ends with the spurt that was mentioned in Bates et al. (1992), and Thatcher's study may help us understand the fact that brain development is a possible candidate for the results found (see also Fischer & Rose, 1994 for an extensive discussion on discontinuities in brain growth, and the relationship with behaviour).

As has been pointed out by Bates et al. (1992), a relationship with grammar exists only on the basis of correlates: “the relationship between synaptogenesis (i.e. rapid acceleration in number of synapses within and across regions of cortex) and language learning is less direct, i.e. an *enabling relation*” (p. 105). A brain growth burst eventually provides more storage capacities, which enhances the speed of the grammatical process. This is supported by three facts. First, the process of word learning is already on its way, and it has consequences for the grammar to be build. That is, if more words are learned, a child may lengthen his sentences to express relationships between words. This is the first step in development which leads to structural patterns (like SVO patterns in English and SOV patterns in Dutch). Second, capacity-related changes in non-linguistic cognition correlate with grammatical changes, which suggest that language and non-linguistic development are timed by a common factor. Thus, it is *timed* by a common factor, and therefore there may be a relation between both, so that correlations may be high, but it does not explain the grammatical changes. Third, language development is not the consequence of a specific and localized language module that guides the transition from first words to grammar. Evidence for this is that children can recover from left-hemisphere damage after the second year of life. The switch from first words to grammar is not the finish of hemisphere specialisation, but an intermediary stage of changing neurological patterns.

It is likely that change is induced by neurological changes, but it is unlikely that neurological changes are responsible, in a one to one scenario, for the development of specific domains of language. Interestingly, in my overview of the catastrophe model, I argued that this implication does not necessarily follow from the premise that the overt use of already present function words (as instantiations of functional categories) depends on gradually increasing performance resources. In fact, if they are present as an underlying or latent structure, catastrophe theory predicts that their emergence in the form of overt speech occurs more or less explosively. In this sense, catastrophe theory is compatible with the assumption that language production is the expression of underlying syntactic rules that are either present at the beginning or emerge at a later stage but are in any case dependent on performance limitations that must be overcome somewhere along the developmental path. However, if the process underlying the acquisition of function words is one of gradually expanding the domain of application, by means of learning and/or imitation of phrases in which function words occur, we expect to find a gradual increase pattern with fluctuations that amount to mere stochastic variance. Recall that, according to catastrophe theory, a continuous decrease of performance limitations does *not* lead to the prediction of continuous increase in the frequency of function words, especially not if the use of function words depends on the presence of latent factors such as underlying syntactic knowledge. This continuous increase of function words is not only depending on performance limitations, but also on the value of the splitting parameter in the Cusp model. If this value is high, the developmental path taken is a discontinuous one.

Increase of memory

Although an increase in the brain volume that is responsible for a growth in memory might be a ‘correlate’ (Bates et al. 1992) of language development, it is not plausible that changes in a specific language domain (for example, functional categories) are caused by specific brain areas. However, the increase of synapses is probably responsible for a change in memory capacity. This could lead to an (sudden) increase of function words. So, Bates Thal and Janowsky found, so to speak, “soft evidence” for a relation between function words and an increase of memory. This soft evidence is a correlate between brain spurts, memory and language development.

An increase of memory or attention (due to, for example, neurological changes) could lead to the increase of function words, but immediately the question “But why function words?” pops up. Hard evidence for the relation between function words and memory is given by Elman (1993). He has shown that the limitation of short term memory can act as a source for increasing complexity. This is what he calls ‘starting small’ (in terms of memory capacity), and it might contribute to the fact that attention is given to specific parts of the environmental input that are developmentally most relevant. For example, due to a limited working memory span, attention is restricted to nouns and verbs in the early stages of child language. This in turn might lead to child language in which function words are temporarily missing. Elman’s theory is an example of a positive instead of a restrictive interpretation of the function of performance limitations. In sum, function *words* are presumably left out in early child speech because of certain (non-)linguistic factors like phonological constraints or memory limitations. This does not rule out, however, that the relevant functional *categories* may be present on an underlying level. Elman’s finding does not explain the rapidity of change either.

The shape and timing of function words in child language: a summary

Two issues constitute the core of this section: the timing and shape of development. I related evidence for both aspects into a converging model of language development. Here I shortly present a model that pays attention to the linguistic and psychological reality of development.

Since the forms of change were the main subject of this study, I will discuss them first, with a special emphasis on the relationship between the data in chapter 5 and the theories on (language) development. However, I argue that the timing of function words is not independent of the shape of change.

The quantitative explanations stretch out over the dynamic systems that were used for the fitting of the empirical data. One reason for jumps might be that a variable (e.g. closed class words) gets support from another grower, which results in closed class words growing faster. The disappearance of a competing relationship (e.g. competition with two-word phrases) may also lead to a sudden increase in growth rate and a new equilibrium. Precursor relationships cause a new grower to start growing. Growers in a system or model are forced to grow because there are *attractors*. Attractors are low energy states, that are not prescribed in a system, i.e. in the language itself, but the result of word combinations, i.e. the process of a developing grammar. The organisation of grammatical features is inherent in the way linguistic theory argues that relationships between syntactic elements are organised. These syntactic features are acquired due to attractors in development.

As it turned out, the different styles, i.e. pronominal and nominal ones, account for individual differences. These differences are found in the rapidity (in terms of age) of the use of, for example, pronouns. Nominal children in the early stages use mainly nouns to refer to other people and things. The differences in style do not account for the difference in speed of development. There is only a difference in timing: pronominal children are earlier (in age) in their acquisition of function words, nominal children will catch up later. Apart from styles, there could be differences between boys and girls. Due to a curious coincidence, all children in the study are *boys*. There are no qualitative differences between girls and boys expected with respect to the development of syntax, and quantitative differences exist only in the timing of the start of the growth of language. With respect to the shape, there is no research that suggests that there are gender differences in the form of *quantitative* development. In short, all children acquire an adult level of linguistic competence. Since there are no qualitative differences between boys and girls, the predictions of the developmental order of a syntactic variable have to be the same for girls and boys.

I discussed the relationship between the data, non-linear theories and linguistic theories. Elsewhere we discerned, at a general level, three types of relationships between a structural theory of development and a growth model (Ruhland et al. 1995): independence (the structural model and the growth model may be completely autonomous, the growth model neither supports, nor refutes the assumptions of the structural model), compatibility (inductive construction that leads to quantitative predictions, and deductive construction, i.e. by choosing its parameters on the basis of notions and considerations supplied by a structural model) and contradiction. A dynamic model needs a structural theory. With this theory structural change is related to quantitative change. For example, if there is an increase in the carrying capacity K, if there are supportive or competitive factors that push a system to a new level (where K itself does not change, but other factors raise or lower the value of K), or competitive factors that avoid a grower to reach a new level with as a result that the system does not structurally change, a linguistic theory explains the relationships between variables and the quantitative change of these variables.

Structural theories and hypotheses fail to make exact predictions about the forms of change. However, they do provide hints about the particular directions in which change might take place. All predictions can be related to a dynamical systems approach. It is plausible that function words follow a path of mutually-related development, which is supported from a linguistic point of view by assuming that if functional categories are not available from the beginning (the reduced competence), these categories develop simultaneously.

This timing of change is explained by a language internal change. I argued that children do not have problems with the comprehension of function words. Furthermore the developing language system is unstable in the third year of life, i.e. in the period when children acquire function words. Developmental stuttering is a well known phenomenon: children have problems in planning their sentences. Given these facts, and the findings of this study, I propose the following model. Children have lengthened their sentences by the time they acquire inflected verbs. These inflected verbs, in a manner of speaking the glue between the acquired lexical categories, release functional categories. If we add to the inflection of verbs that in the same age span phonological constraints gradually disappear and that, probably due to changes in the brain (i.e. synaptogenesis), memory increases so that function words may appear, the converging evidence helps to understand the timing and shape of function words in development.

In sum, linguistic theories provide qualitative explanations for the development of function words. However, these explanations only concern the relative moment or order of development, not the speed or shape of development. That is, the development of function words is explained in terms of order, not in terms of suddenness. A dynamic systems explanation is a resources dependent explanation. The two sorts of explanation combined are complementary or even overlapping, help to explain the speed of change.

6.3 Non-linear models of development

In this section, I discuss the relationship between several non-linear models. These models help us to untangle two problems. First, there is an account for the shape of change, both continuous and discontinuous. Second, there is a solution for the innateness problem. The first issue on shape is discussed in Changes in language development, the second issue is discussed in the section on Connectionism. The goal is to show that the application of non-linear models reaches further than a discussion on the shape of change, and that we ultimately come to a model of language development that is not build on rules but on relative simple inductive reasoning.

Changes in language development: two brotherly theories

Dynamic systems theory

It is assumed that developmental processes guide the acquisition of language structures. Whether they remain the same over time (age) in the period they are acquired is questionable, since language development is not local, it does not take place within the same age and module (in terms of sorts of linguistic structures that are acquired) domain. Rather, development is a mixture of processes that stretch out over the local mode. Suppose discontinuous changes occur in language development. These changes may, in the course of development, be taken over by a continuous (e.g. learning) model. In that case all energy, attention, motivation, etc. can be spend on other variables and processes.

In the following example, this idea is worked out for child language. Suppose a child is learning to inflect verbs. First, he needs to know what verbs are. A simple rule may help the child to select the right words that are verbs. Since this is not a study on the basic problems of mechanisms, it is taken for granted that a child somehow knows what verbs are. Once he has mastered the inflection rule, he is able to switch his attention to other language variables. Verbs undergo a process of automatism. From that moment on the child may spend time and energy on acquiring other language variables. It is assumed that vocabulary, sentence lengthening, and the inflection rules for verbs (amongst other variables, and other modules like semantics or pragmatics) are precursors of functional categories (Ruhland et al. 1995, Cats & Ruhland, in press). Once verbs (the last 'acquisitional item' before function words emerge) set free other functional categories, verb inflection no longer consumes energy, in terms of, for example, attention, effort, memory load, to process the knowledge of inflection. All time and energy may be spend on the development of function words. This development of function words is, so to say, the result of selforganising structures. The interpretation of the language the child hears is switched from inflection rules to rules that address other syntactic slots. The local model of language development is directed to another variable, that is supported by control variables. The developmental model is assimilated into a dynamic systems model.

Observational data are in general very well suited for applying dynamic growth theory, which describes the transient processes of change towards an eventual equilibrium level. Observational studies of the kind reported in Ruhland and van Geert (1998a) will in general not suffice to really prove the existence of a discontinuity in the underlying equilibrium level. Some of the flags are indicative of the occurrence of a cusp catastrophe, but require the experimental manipulation of relevant control variables. In a dynamic growth model the variables are also control variables of each other. The problem with language, however, is that the control variables are not yet known, or are too general (like the length of the utterance), or are local and temporal (like related linguistic rules or properties whose emergence supports that of new principles or rules). The notion of discontinuities in equilibrium level, however, is naturally related to structural theories that view language development as the emergence of rule-governed structures, of parameter setting etc.

One of the advantages of our model is that it is possible to make predictions on the process that is covered by the model. If we change values, for example, the growth parameter of one grower, the result is most likely that not only that specific grower is changing (e.g. the speed of change), but since it is assumed that growers are coupled (i.e. the growers use the same resources), the change of one parameter causes the whole system to change. Of course, growers use only some of the sources at any moment in development. For example, when a child has mastered the phonology of his language, which, for instance, influences the development of a vocabulary, the phonology is internalised, which means that phonology no longer relies on memory. This change of individual growers may also depend on a change of the parameters in the equations. Basically, the value or strength of the parameters in a model may either stay the same over age, change at every step in time, or change due to temporary changes in other variables or resources. Which one is most likely in development? If parameters remain constant over age, this means that resources remain constant over time. This is not likely: memory, for instance, increases gradually over age (Case 1985). If parameters that influence equilibria change constantly over age, we would not find equilibria. Most likely, therefore, is the assumption that parameters change locally and temporarily. Age and resource specific models, i.e. models with parameters that change due to a change in one or more resources, have yet to be developed.

A quantification of language development uncovers the process of acquisition. Fitting the data with non-linear equations is one step of applying dynamic systems theory to development, but there is another use of the equations. This involves the building of a model. This model (Cats & Ruhland, in press) is a dynamic systems model build of non-linear equations. A non-linear model makes it possible to predict dependent variables in relationship with each other. There is a mutual relationship between growers, since they rely on the same psychological resources like memory, attention, etc. The model describes individual trajectories, which means that situational, temporal and individual variation is part of the model. Furthermore, these models describe a process, not a product, which is a fundamental insight in change as it occurs in development. Basic characteristics are that in these models variables are mutual and non-linearly related. In a mathematical sense, the equations are iterative, that is, the present state of a system is depending on the previous state, and the output of the previous state is the input of the present state. Intriguingly, jumps, fluctuations and regressions are intrinsic to the system. That is, they are not part of the equations that monitor the process in the model.

A theory is needed to link the equations to a process. Any random series of numbers may be fitted with an (non-)linear equation. However, a model with only equations is just a fitting procedure. The prediction are empty. A structural or functional model prescribes, first, the variables in a model, and, second, the relationships between the variables. Furthermore, since there are several variables and parameters in the non-linear equations, a model of connected growers would lead to an enormous amount of the parameters. A reduction of this amount of parameters is needed because an overload of parameters will eventually fit any series of numbers. In other words, every data point (empirically or randomly chosen) is fitted by an individual parameter. This is in fact not a model (i.e. a reduction of reality), but reformulation of reality in terms of parameters.

The construction of a model needs to consider language impairments, e.g. deafness. If these impairments yield different timing or shape patterns of development, and if the dynamic systems model is correct, these patterns must be part of the model, under the assumption that the basic structural patterns do not differ from normal children. However, longitudinal studies with dense time series of language impaired children are scarce, so conclusions are hard to reach, but the main thinking is research on language impaired children could provide extra evidence that a dynamic systems model is correct, i.e. that development depends on other growers. New research might reveal some interesting views on the development of, for example, language delayed or impaired children. The evidence of normal and impaired children may converge to a model of change, which correctly reflects language development.

Lastly, the model may be so strong that any data set may be explained with the model. A random model with any choice of the variables and their parameters is so strong that there is an overkill of parameters. To avoid this problem, one must also try to model random series. If the model is still capable of producing good fits, the model is simply too strong to be applied to development. In future research (Ruhland, 1998b) building models and testing these models against random time series will elicit the power of these models.

Catastrophe theory

The choice for catastrophe theory and its flags as the theory and model for development has its limitations. Although development is best fitted by the cusp, one might object that a Catastrophe model is not a real model for development since it is a local model. A local model explains changes on a very small 'window' of time, e.g. age. Being a local model, only temporary changes are explained. Since development takes place over time, a model of language development has to extend beyond the local 'mode'. Catastrophe theory models phase transitions. The strong point of this theory are the criteria, i.e. the flags, that must show up together. The density of the data need a minimum since three points studies reveal a linear increase (Ruhland, 1998b), but since a rapid increase and multimodality is found using two weekly recordings, catastrophe theory and its flags is applicable to the data in this study.

Elsewhere I argued that both observational studies and experiments (i.e. a structured form of observation) are required to reach the heart of language development (Ruhland, 1991). Catastrophe theory quite clearly shows that this is a correct observation. The stability of a system can only be tested by using experiments. However, the first step in the application of non-linear models was to use longitudinal observational data. Now we have a better view on processes in language development we can start to use the technical form of observation.

The exponent or power of the non-linear equations of a dynamic systems model acts upon the shape of the equation. If the power of an equation is close to 2, the equation shows a form of change that resembles the sudden jump in Catastrophe Theory. This theory is useful for the detection of phase transitions. However, not all (also non-language) growers take the shape of a rapid increase or even discontinuous change. To meet other developmental paths that have to be fitted by a restricted class of equations, dynamic systems theory is complementary to Catastrophe Theory.

The non-linear models that have been discussed so far cannot explain the emergence of language structures, because these models describe change in development. Connectionism shows us how account for the innateness problem.

Connectionism

Connectionist models have been applied to language development by Elman et al. (1996). Their book, *Rethinking innateness*, elaborates the question to what extent innateness is needed to explain the acquisition of language (from phonology to syntax). Dynamic systems models are useful to fit developmental paths, as this study shows. However, dynamic models can not be used to study the genesis of behaviour. Recently, the discussion on innateness and new behaviour has been taken up by Elman and others (Elman, 1993, Elman et al., 1996). But what is connectionism? Connectionism is a family of computational principles. Simple Hebbian rules (i.e. learning rule that specify how much the weight of the connection between two units, for example, nodes in a neural network, should be increased or decreased in proportion to the product of their activation; this activation is also closely connected to the input that, for instance, a neural network receives) lead to learning that is not the linear result of those rules. Connectionist models tend to make use of constraint satisfaction, distributed representations, learning on the basis of local information only, etc., but there are plenty of connectionist models that fail to follow one or another of the characteristic principles.

Connectionism can be used to implement any theory. In the discussion of these models, for example, a model of the past tense, it is not the question whether some connectionist model can handle the past tense, but rather, what would a connectionist model that did handle the past tense have to look like. For example, in English people appear to freely generalize the -ed morpheme to words with unfamiliar sounds.

6.3. Rick webbed today.

Likewise the Dutch use the -te/-de morpheme to establish the past tense of a new verb. Standard models tend to have trouble accounting for this phenomenon without the assumption of innateness.

The difference between connectionism and dynamic systems models is that the first aims at explaining the (epi)genesis of behaviour, while the latter are needed to fit growth curves to find out about the shape of change. A second difference is that connectionist models are learning models. This means that such a model receives any input, and on the basis of the input, the model renders some output, whereas dynamic systems models try to explain the form of change and interactions between variables. Both models may be at their best if we combine them. First, we have an explanation and a description of the process. Discontinuity is demonstrated by Catastrophe theory. It is not a vague theory that pours old concepts from, for example, a Piagetian framework into a formal theory, and it is applied to development using indicators for discontinuous change. Dynamic systems theories describes continuous, but non-linear change. This non-linear change covers the dependence on resources, the growers that are connected, and modelling of humps and bumps as basic elements of development. However, these two models do not explain the origin of development. Second, we have a group of models that can explain the origin of development without the assumption of innate structures. Connectionist models get us out of the discussion on whether or not innateness is needed. The charm of both sorts of models, i.e. process and origin models, is that we do not need a deus ex machina or a homonculus to explain where our knowledge comes from, neither do we need to assume linear change as **the** basic form of development. What we do need in both cases are the right models that can explain the structural component.

The assumption of linear change and innateness is too simple to capture the richness of a language, and of a process. Selforganising patterns, as a result of the language itself, emerge from processes on a ontogenetic scale. Due to, for example, memory, phonological and rhythmic patterns, and syntactic ordering, the child starts to play with language, and at the same time, he discovers patterns of his language. Non-linear models explain development without the use of prescribed states or structures. This, however, does not mean that we do not need the linguistic theories that I discussed in chapter 2. These theories, which ever one is chosen, give a thorough underpinning of the variables that should be studied in language development, and these theories can explain order and structure in development. Crucial in a design of non-linear modelling is that linguistic entities and psychological processes meet.

6.4 'Houston, Tranquillity base here: the Eagle has landed': final remarks

The non-linear fitting of data and the building of models are necessary to uncover growth and development. Structural models provide the qualitative underpinning of the research domain to which non-linear models were applied to. Of course, non-linear and structural models stress different issues. Furthermore, the flags from catastrophe theory have no definite, conclusive value. They are indicators of discontinuities, i.e they make it plausible that development is discontinuous. The changes found and presented constitute a small part of the human time scale, namely the change from the prefunctional category stage to the functional category stage. It is not possible, and also not desirable, to generalize over other age periods. In addition, there may be a simple reason why no catastrophic change was found: children have the option of employing different styles to their personal development. This, however, does not follow directly from studies on the acquisition of function words, since there is no compelling reason to believe that two styles of language make a catastrophic change impossible. In fact, any change is possible, since there is no obligation of changing discontinuously, or even non-linearly. The data in this study show the empirical foundation of this individual variability.

The evidence presented in this chapter converges into a framework, where timing and shape of change contribute to the development of syntax. This developmental, mathematical psycholinguistic approach reveals the two important issues in the study of language development: structure and process are like brother and sister. The claim of transitional change (i.e. the flags of catastrophe theory) gets stronger if additional catastrophic flags are found. An experimental setting is needed to find (some of) the other flags, such as critical slowing down. This setting makes it possible to perturb a system. Perturbations test the stability of a system undergoing change. The stability is better if the speed of slowing down after a perturbation is high. Furthermore, hysteresis, which is a strong explanatory flag, requires an experimental setting (see also chapter 3 for an example of hysteresis in an experiment). The discovery of the control parameter(s) makes it possible to manipulate a system. Words per session or the size of working memory might be appropriate candidates. I presented a count of words per session, but the problem is that this variable does not change linearly. A rapid or stepwise increase in the control variable (e.g. in Peter's case) is problematic: this sudden increase makes it hard to prove that the change in a dependent variable is caused by a small change in de independent (i.e. control) variable. More catastrophe flags are needed, especially those that call for a manipulation of one or more control variables. These control variables have to be manipulated in a linearly way to find, for instance, divergence. Further research, with an experimental setting, is needed to prove the existence of other flags.

Methodologically, this study could be extended to other elementary issues of development. *High density recordings* (e.g. weekly or even daily ones) might prove that fluctuations and variation are characteristic for daily or weekly rhythms. In a pilot study, we collected daily recordings. The analyses of these recordings revealed that two weekly fluctuations can also be found on a daily scale, but these fluctuations appear to be less extreme (i.e. a smaller amplitude) and with a different frequency. In other words, there seems to be stability on a day by day basis (De Goede & Ruhland, 1998). De Weert and van Geert (1997) demonstrate that this daily and weekly variability also exists in social-emotional development. A *data collection* needs the correct time index. The lack of recordings in between two weekly recordings may be problematic, since there may be fluctuations between the sessions. However, daily recordings have two objections. First, there are practical objections. A daily recording schedule is labourious. Second, time series require a certain density to uncover developmental paths. Traditionally, this density has been treated in a rather simple way, namely by using three data points with an interval of several months. Three points, e.g. one occurring two months before the transition, one in the transition period and the third one two months after the transition, make the development of a variable look more or less linear (Ruhland, 1998b). Recording every two weeks seemed to cover the development of syntax most adequately. The present method is meant as a heuristic exploration of the efficacy of the method. Empirical evidence, however, is not yet available. In our model (Cats & Ruhland, in press) an increase of the number of recordings beyond a certain frequency did not lead to a better fit, i.e. there is no improvement with simulations using a one week interval. That is, the simulation showed that with the use of random numbers and weekly intervals there is no significant increase in the explanation of the paths found (Cats & Ruhland, in press).

The *length of the recording* was one hour (cf. Nelson, 1973). The question is: does the increase of recording time yield other insights? The answer is unclear. One reasoning would be that an increase of recording time probably will only lead to quantitative differences. The absolute numbers of the use of a variable changes during the increased recording time, not the overall qualitative shape of change. Of course, the recording time in the sessions should have a minimum. Since half hour sessions (used as an index of anomalous variance) show considerable differences in at least some of the children, recordings shorter than an hour are dubious. Research using longer recording sessions still awaits the rigours of (statistical) testing. The importance of the length of recording is at its height in a transition period, i.e. when there is no stable equilibrium. We are now performing a day by day study to find out more about the variability on the day level. *Daily variability* may be sensitive to the moment of the day (e.g. morning vs. afternoon). Smaller scale analyses, on a day and week level, are now interesting, since the general and individual patterns are known.

The *choice* of function words leads to the question whether function words change qualitatively during development. This is an important question in a study on development and change. There is some evidence that determinative pronouns like *this* and *that* change during development: children use these words with pointing in the early stages of development. However, within the age range of the sudden change (in the third year of life), the local system of functional categories and function words is assumed to remain constant. Evidence that they do not change is that function words have a constant place in the utterances of a child. Furthermore, there is a constancy in the input: function words have a constant place in the parental sentences. The frequency of function words of the parents, and their use of function words is high. In short, there is no reason to assume that these words qualitatively. The study of quantitative language development must, in order to prevent such qualitative change to happen, rely on linguistic studies that reveal through experiments whether qualitative change is taking place.

The end

One of the biggest achievements of mankind in this century was the first man of the moon. It was more than a step forward, even for mankind. The techniques and procedures, the use of computers, and the wish to go to the natural satellite that circles the earth, led to the first footprints on the moon. In a similar fashion, but on a much smaller scale, this study seized the opportunity, within the field of human development, of attacking the problem of how to analyse individual paths of development. I think it is fair to say that this exploration of new worlds in language development research has been successful in terms of the landing of a new method on child language. In comparable terms of the landing on the moon, this study revealed that linear and unimodal analyses of development, or structural linguistic analyses of language development alone, are methods that are too narrow, too restricted to describe and explain all aspects of language development. If we apply these new methods to language development, we will come up with new insights, and with panoramic views, just like the astronauts of Apollo 11 who successfully landed on the moon and went live on television in July 1969. I quote: “We have a beautiful view of the earth here, it is absolutely fantastic.” It is the difference between Apollo 11 and 13 (which was the Apollo mission that got into big problems after one of the oxygen tanks had been blown up), it is the difference between a successful landing and a near-catastrophe. After all, it’s numbers that make the difference.

Voyages to other planets and other lunar systems are still undertaken to seek new life forms and to boldly go where no man has ever gone before. For example, Mars has been visited in summer 1997. In psychology, researchers have boldly gone where no researcher has gone before with introduction of new methods for the study of development of linguistic structures. This view is brought into focus by non-linear models. The reasoning is build on the combination of disciplines, i.e. the use of complementary approaches. The data collection consists of intensive recordings of child language over a long period. Frequency counts were used, and the results support non-linear, continuous development. I found complementary evidence for a reorganisation in language development, but I do not intend to claim that this is the end of non-linear analyses. This study is the start of more interdisciplinary studies using new techniques like non-linear models and computer assisted analysis. Especially the link between the origin and shape of behaviour, structural models, quantitative analyses, and brain structures (as the origin of behaviour) need more study. Furthermore, the question is whether other language variables than function words undergo the same change. The acquisition of a rule like verb finiteness (Wijnen, 1996), for example, is a gradual process. Future research should elucidate rapid or transitional change in other variables or on other linguistic levels (e.g. on a phonological or semantic level). Such evidence has been stated verbally by Fikkert (1992), who found a more or less sudden change in the mastering of stress patterns. The question is whether this change is also sudden in terms of a growth model or in a catastrophe model. The answer to this question and other questions on changes in language development lies ahead, and it is a fascinating and promising challenge for non-linear models.

This study is one out of a growing body of studies that hopefully mark a changing tide. Not only because it has an interdisciplinary character using linguistics and psychology, but also because it uses new techniques of non-linear modelling, and dense time series to elicit the shape of change.

References

Atkinson, M. (1986) 'Learning and Models of Development.' In P. van Geert (ed.) *Theory building in developmental psychology*. Amsterdam [etc.]: North-Holland.

Bates, E., Bretherton, I. & Snyder, Lynn S. (1988). *From first words to grammar.* New York, NY: Cambridge University Press.

Bates, E., Thal, D. & Janowsky, J.S. (1992) 'Early language development and its neural correlates.' In S.J. Segalowitz & I. Rapin (eds.) *Handbook of neuropsychology (Vol. 7: Child Neuropsychology).* Amsterdam [etc.]: Elsevier.

Behrens, H. (1993) *Temporal reference in German child language: form and function of early verb use.* Zutphen: Koninklijke Wöhrman.

Bloom, L., Lightbown, P.M. & Hood, L. (1975) 'Structure and variation in child language.' *Monographs of the Society for Research in Child Development* **40**, 2, No. 160.

Bohn, O.-S. & Flege, J.E. (1993) 'Perceptual switching in Spanish / English bilinguals.' *Journal of Phonetics* **21**, 267-290.

Bol, G.W. (1996) 'Optional Subjects in Dutch Child Language'. In Ch. Koster-Kromhout & F. Wijnen (eds.) *Proceedings of the Groningen Assembly on Language Acquisition*. Held at the University of Groningen, 7 - 9 September 1995. Centre for Language and Cognition Groningen: Groningen.

Boom, J. (1993) *The concept developmental stage*. Nijmegen: Universiteitsdrukkerij Nijmegen.

Boom, J., Gerlagh, G. & Hartog, A. den (1997). 'The waterlevel task and hysteresis.' Poster presented at the Dutch Developmental Psychology congress, 17-18 April 1997, Leusden, The Netherlands.

Bower, T. (1983) *Development in infancy.* San Francisco: Freeman.

Bowerman, M.F. (1983) 'The 'No negative evidence' problem: how do children avoid constructing an overly general grammar?' *Papers and reports on child language development* **22**, 23-35.

Braine, M.D.S. (1992) 'What sort of innate structure is needed to 'bootstrap' into syntax?' *Cognition* **45**, 77-100.

Braine, M.D.S. (1994) 'Is nativism sufficient?' *Journal of child language* **21**, 9-31.

Brainerd, C.J. (1978) 'The stage question in cognitive-developmental theory.' *The behavioral and brain sciences* **2**, 173-213.

Brandenburg, G. (1915) 'The language of a three year old.' *Pedagogical seminary* **22**, 89-120.

Brown, R. (1973) *A first language: the early stages.* London, Allen & Unwin.

Brown, R. (1979) 'Telegraphic speech.' In V. Lee (ed.) *Language development.* London: Open University Press.

Cake (1996) 'The distance'. On *Fashion Nugget.* [CD] Capricorn Records.

Camaioni, L. (1989) 'The role of social interaction in the transition from communication to language.' In A. de Ribaupierre (ed.) *Transition mechanisms in child development: the longitudinal perspective.* Cambridge [etc.]: Cambridge University Press.

Case, R. (1992) *The mind's staircase: exploring the conceptual underpinnings of children's thought and knowledge*. Hillsdale N.J.: Erlbaum.

Cats, M. & Ruhland, R. (In press) 'Simulatie van taalontwikkeling.' [Simulation of language development] To appear in *Tijdschrift voor Informatica en Maatschappij* (TIM).

Cazden, C.B. & Brown, R. (1975) 'The early development of the mother tongue.' In A. Bar-Adon & W.F. Leopold (eds.) *Child language: a book of readings.* Englewood Cliffs, N.J.: Prentice-Hall.

Chabon, S.S., Kent-Udolf, L. & Egolf, D.B. (1982) ‘The temporal reliability of Brown’s mean length of utterance (MLU-M) measure with Post-Stage V children.’ *Journal of speech and hearing research* **25**, 124-128.

Chomsky, N. (1986) *Barriers.* Cambridge, Mass. [etc.]: MIT Press.

Clahsen, H., Eisenbeiss, S. & Penke, M. (1994) ‘Underspecification and lexical learning in early child grammars.’ *Essex research reports in lingustics* **4**, 1-36.

Clahsen, H., Penke, M. & Parodi, T. (1994) ‘Functional categories in early child German.’ *Language acquisition,* **3** (4), 395 - 429.

Cobb, L. (1978) ‘Stochastic catastrophe models and multimodal distributions.’ *Behavioral science* **23**, 360-374.

Cobb, L. (1981) ‘Parameter estimation for the cusp catastrophe model.’ *Behavioral science* **26**, 75 - 78.

Connell, J.P. & Furman, W. (1984) ‘The study of transitions: conceptual and methodological issues.’ In R.N. Emde & R.J. Harmon (eds.) *Continuities and discontinuities in development.* New York: Plenum.

Davison, J., Koch, H.W.Jr., Lowry, H. (Producers), & Abrahams, J., Zucker, D., Zucker, J. (Directors). (1980). *Airplane!* [Film]. Available from Paramount Pictures Corporation.

Donahue, M. (1986) ‘Phonological constraints on the emergence of two-word utterances.’ *Journal of child language* **13**, 209-218.

Dore, J., Franklin, M.B., Miller, R.T. & Ramer, A.L.H. (1975) ‘Transitional phenomena in early language acquisition.’ *Journal of child language* **3**, 13-28.

Dromi, E. (1986) ‘The one-word period as a stage in language development: quantitative and qualitative accounts.’ In I. Levin (ed.) *Stage and structure: reopening the debate.* Norwood New Jersey: Ablex.

Eimas, P. (1985) ‘The perception of speech in early infancy’. *Scientific American* **252**, 45-63.

Elbers, L. & Wijnen, F. (1992) ‘Effort, Production Skill, and Language Learning.’ In A. Ferguson-Charles, L. Menn & C. Stoel-Gammon (eds.) *Phonological Development: Models, Research, Implications.* Timonium, MD: York.

Elman, J.L. (1993) ‘Learning and the development in neural networks: the importance of starting small.’ *Cognition* **48** (1), 71 - 99.

Elman, J.L., Bates, E.A., Johnson, M.H., Karmiloff-Smith, A., Parisi, D. & Plunkett, K. (1996) *Rethinking innateness: a connectionist perspective on development.* Cambridge, MA: MIT Press.

Emde, R.N. & Harmon, R.J. (1984a) *Continuities and discontinuities in development.* New York: Plenum.

Emde, R.N. & Harmon, R.J. (1984b) ‘Entering a new era in the search for developmental continiuities.’ In R.N. Emde & R.J. Harmon (eds.) *Continuities and discontinuities in development.* New York: Plenum.

Everitt, B.S. & Hand, D.J. (1981). *Finite Mixture Distributions*. London: Chapman and Hall.

Felix, S. (1987) *Cognition and language growth.* Dordrecht [etc.]: Foris.

Felix, S.W. (1992) ‘Language acquisition as a maturational process.’ In J. Weissenborn, H. Goodluck & T. Roeper (eds.) *Theoretical issues in language acquisition: continuity and change in development.* Hillsdale, N.J. [etc]: Erlbaum.

Fenson, L., Dale, P.S., Reznick, J.S. & Bates, E. (1994) ‘Variability in early communicative development.’ *Monographs of the Society for Reseach in Child Development* **242** (59, 5). Chicago: University of Chicago Press.

Ferdinand, A. (1996). *The development of functional categories: the acquisition of the subject in French.* The Hague: Holland Academic Graphics.

Fikkert, P. (1992) ‘The acquisition of Dutch stress.’ In M. Verrips & F. Wijnen (eds.) *Amsterdam Series in Child Language Development:the acquisition of Dutch* **5**, 21-35.

Fischer, K.W. (1983a) ‘Developmental levels as periods of discontinuity.’ In K.W. Fischer (ed.) *Levels and transitions in children's development.* San Francisco: Jossey-Bass.

Fischer, K.W. (ed.) (1983b) *Levels and transitions in children's development.* San Francisco: Jossey-Bass.

Fischer, K.W., Pipp, S.L. & Bullock, D. (1984) 'Detecting developmental discontinuities.' In R.N. Emde & R.J. Harmon (eds.) *Continuities and discontinuities in development.* New York: Plenum.

Fischer, K.W. & Silvern, L. (1985) 'Stages and individual differences in cognitive development.' *Annual Review of Psychology* **36**, 613 - 648.

Fischer, K.W. & S.P. Rose (1994) 'Dynamic development of coordination of components in brain and behavior; a framework for theory and research.' In G. Dawson & K.W. Fischer (eds.) *Human behvaior and the developing brain.* New York [etc.]: The Guilford Press.

Fischer, K.W. & Granott, N. (1995) 'Beyond one-dimensional change: parallel, concurrent, socially distributed processes in learning and development.' *Human Development* **38**, 302-214.

Fodor, J.A. (1983) *The modularity of mind.* Cambridge, Mass. [etc.]: MIT Press.

Ford, D.H & Lerner, R.M. (1992) *Developmental systems theory: an integrative approach.* London [etc.]: Sage.

Fogel, A. (1992) 'Movement and communication in human infancy: the social dynamics of development.' *Human movement science* **11**, 387-423.

Gerken, L.A., Landau, B. & Remez, R.E. (1990). 'Function morphemes in young children's speech perception and production.' *Developmental psychology* **26** (2), 204-216.

Gerken, L.A. & McIntosh, B.J. (1993) 'Interplay of Function Morphemes and Prosody in Early Language.' *Developmental Psychology* **29** (3), 448-457.

Gibson, K.R. & Petersen, A.C. (1991) *Brain Maturation and Cognitive Growth: comparative and Cross-Cultural Perspectives.* New York: Aldine de Gruyter.

Gibson, E. & Wexler, K. (1994) 'Triggers.' *Linguistic Inquiry* **25**, 407-454.

Gilbers, D.G. (1992) *Phonological networks: a theory of segment representation.* PhD. Dissertation, University of Groningen. (Grodil series, no. 3).

Gilmore, R. (1981) *Catastrophe theory for scientists and engineers.* New York: John Wiley & Sons.

Glaserfield, E. von & Kelly, M.F. (1982) 'On the concepts of period, phase, stage, and level.' *Human development* **25**, 152-160.

Gleick, J. (1989) *Chaos.* Amsterdam: Contact.

Gleitman, L.R. & Wanner, E. (1988) 'Current issues in language learning.' In M.H. Bornstein & M.E. Lamb (eds.) *Perceptual, cognitive and linguistic development.* Hove [etc.]: Erlbaum.

Goede, D. de & Ruhland, R. (1998) 'Daily rythms in language development.' Manuscript in preparation.

Goldin-Meadow, S., Alibali, M.W. & Church, R.B. (1993). 'Transitions in concept acquisitionm: using the hand to read the mind.' *Psychological review* **100** (2), 279-297.

Golinkoff, R.M. (ed.) (1983) *The transition from prelinguistic to linguistic communication.* Hillsdale, NJ.: Lawrence Erlbaum Associates.

Golinkoff, R.M. & Hirsh-Pasek, K. (1990) 'Let the mute speak: what infants can tell us about language acquisition.' *Merrill-Palmer quarterly* **36** (1), 67-92.

Goodluck, H. (1991) *Language acquisition: a linguistic introduction.* Oxford [etc.]: Blackwell.

De Graaf, J.W. (1997) *Relating new to old.* Unpublished dissertation, University of Groningen.

De Graaf, J.W. & Breeuwsma, G. (1996). 'The advantage of being illiterate: a constructivistic theory of language acquistion in accounting for the critical period phenomenon.' Unpublished manuscript.

Gunnar, M.R. & Thelen, E. (eds.) (1989) *Systems and development.* Hillsdale, N.J. [etc.]: Erlbaum.

Hartelman, P.A.I. (1996) 'Cusp Catastrophe Estimation.' Internal manuscript, Department of Psychology, University of Amsterdam.

Higginson, R. & Macwhinney, B. (ed.) (1991) *Childes/BIB: An annotated bibliography of child language and language disorders.* Hillsdale, N.J.: Lawrence Erlbaum Associates. [Supplement 1994]

Hoekstra, T. & Jordens, P. (1994) ‘From adjunt to head.’ In T. Hoekstra & B. Schwarz (eds.) *Language acquistion studies in generative grammar: papers in honor of Kenneth Wexler.* Amsterdam: John Benjamins.

Hoppe-Graf, S. (1989) ‘The study of transitions in development: potentialities of the longitudinal approach.’ In A. de Ribaupierre (ed.) *Transition mechanisms in child development: the longitudinal perspective.* Cambridge [etc.]: Cambridge University Press.

Hothouse Flowers (1993). ‘Your nature.’ On *Songs from the rain.* [CD]. Polygrm Records.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. & Lyons, T. (1991) ‘Early vocabulary growth: relation to language input and gender.’ *Developmental Psychology* **27**, 236-248.

Hyams, N.M. (1986) *Language acquisition and the theory of parameters.* Dordrecht [etc.]: Reidel.

Hyams, D. (1996) *CurveExpert* [Computer Software].

Ingram, D. (1989) *First language acquisition: method, description and explanation.* Cambridge: University Press.

Jackson, J. (1979). ‘It’s different for girls.’ On *I’m the man* [CD]. A & M Records.

Karmiloff-Smith, A. (1992) *Beyond modularity: a developmental perspective on cognitive science*. Cambridge [Mass.]: MIT Press.

Kaufman, S.A. (1993) *The origins of order: self-organization and selection in evolution.* Oxford [etc.]: Oxford University Press.

Koster, J. (1976) ‘Dutch as a SOV language.’ *Linguistic analysis* **1**, 111-136.

Koster, C. (1993) *Errors in anaphora acquisition.* Utrecht: Led.

Levelt, W.J.M. (1989) *Speaking: from intention to articulation.* Cambridge, Mass. [etc.]: MIT Press.

Maccoby, E.E. (1988) ‘Gender as a social category.’ *Developmental Psychology* **24** (6), 755-765.

MacWhinney, B., Leinbach, J., Tabaran, R. & MacDonald, J. (1989) ‘Language learning: Cues or rules?’ *Journal of Memory and Language* **28**, 255-277.

MacWhinney, B. & Snow, C.E. (1990) ‘The child language data exchange system.’ *Journal of child language* **12**, 271-296.

Macwhinney, B. (1991) *The CHILDES Project: Tools for analyzing Talk.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Macwhinney, B. (1997) ‘Rethinking the logical problem of language acquisition.’ Unpublished manuscript found on the Internet address Http://psyscope.psy.cmu.edu/local/Brian/papers.html.

Magestic [Computer Software]. (1996). Dallas, The United States: Logix Consulting.

Manic Street Preachers (1996). ‘Everything must go.’ On *Everything must go* [CD]. Columbia Records.

McCune, L. (1992) ‘First words: a dynamic systems view.’ In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.) *Phonological development: models, research, implications.* Timomium, Maryland: York Press.

Meisel, J.M. (ed.) (1992) *The acquisition of verb placement, functional categories and V2 phenomena in language development.* Dordrecht, Kluwer.

Miller, P.H. (1993) *Theories of developmental psychology*. New York, Freedman.

Mussen, P.H. (ed.) (1983) *Handbook of Child psychology.* Vol. III: Cognitive development. New York, Wiley.

Nelson, K. (1973) ‘Structure and strategy in learning to talk.’ *Monographs of the Society for Research in Child Development* **38** (1-2, 149).

Nelson, K. (1975) ‘The nominal shift in semantic-syntactic development.’ *Cognitive Psychology* **7**, 461-479.

Oxford Concise Dictionary of current English, The (1995). Edited by Della Thompson. Oxford: Clarendon Press.

O’Grady, W. (1997) *Syntactic development.* Chicago [etc.]: UCP.

Peters, E.E. (1991) *Chaos and order in the capital markets: a new view of cycles, prices, and market. volatility*. New York [etc.]: Wiley.

Petrtic, P. & Tweney, R. (1977) ‘Does comprehension precede production? The development of children’s responses to telegraphic sentences of varying grammatical adequacy.’ *Journal of Child Language* **4**, 56-73.

Plato (1949) *Meno* Published by H.D. Verdam. Groningen [etc.]: Wolters.

Powers, S.M. (1996) *The Growth of the Phrase Markers: Evidence from Subjects*. Doctoral thesis, University of Maryland.

Radford, A. (1988a) ‘Small children small clauses.’ *Transactions of the Philological society* **86** (1), 1-43.

Radford, A. (1988b) *Transformational grammar: a first course.* Cambridge [etc.]: Cambridge University Press.

Radford, A. (1990a) ‘The syntax of nominal arguments in early child English.’ *Language Acquisition* **1** (3), 195-223.

Radford, A. (1990b) *Syntactic theory and the acquisition of English syntax: the nature of early child grammars of English*. Oxford [etc.]: Blackwell.

Radford, A. (1992) ‘The acquisition of the morphosyntax of finite verbs in English.’ In J.M. Meisel (ed.) *The acquisition of verb placement, functional categories and V2 phenomena in language development.* Dordrecht: Kluwer.

REM (1986). ‘Begin the begin.’ On *Lifes Rich Pageant* [CD]. IRS Records.

Richards, B.J. (1990) *Language development and individual differences: a study of auxiliary verb learning.* Cambridge [etc.]: Cambridge University Press.

Riegel, K.F. (1975) ‘Toward a dialectical theory of development.’ *Human development* **18**, 50-64.

Roberts, K. (1983) ‘Comprehension and production of word order in Stage I.’ *Child Development* **54**, 2, 443-449.

Ruhland, R. (1991) ‘De acquistie van ‘hoeven’.’ *Tabu* **21** (3), 97-116.

Ruhland, R., Wijnen, F. & van Geert, P. (1995) ‘An exploration into the application of dynamic systems modelling to language acquisition.’ In F. Wijnen & M. Verrips (eds.) *Amsterdam Series in Child Language Development: approaches to parameter setting* 4, 107 - 134.

Ruhland, R. (1996) ‘Frequency analysis in language development.’ In R. Jonkers, E. Kaan & A. Wiegel (eds.) *Language and Cognition 5 (Yearbook 1995 of the research group for Theoretical and Experimental Linguistics of the University of Groningen)*. University of Groningen, Dept. of Linguistics.

Ruhland, R. (1998a) ‘Closed class words in motion.’ To appear in the proceedings of the KNAW-seminar *Non-linear approaches to development*.

Ruhland, R. (1998b) ‘Theory building depends on data collection.’ Manuscript in preparation.

Ruhland, R. & van Geert, P. (1998a) ‘Jumping into syntax: transitions in the development of closed class words.’ In press. *British journal of developmental psychology*.

Ruhland, R. & van Geert, P. (1998b) ‘Building models of language development: a non-linear approach.’ Manuscript in preparation.

Runrig (1993) ‘Ard.’ On *Amazing things.* [CD]. Chrysalis Records.

Schlesinger, I.M. (1971) ‘Production of utterances and language acquisition.’ In D.I. Slobin (ed.) *The ontogenesis of grammar.* New York: Academic Press.

Serbin, L. A., Sprafkin, C., Elman, M., & Doyle, A.B. (1984). ‘The early development of sex differentiated patterns of social influence.’ *Canadian journal of social science* **14**, 350-363.

Sherman, R.M. & Sherman, R.B. (1967). ‘The bare necessities.’ On *The Jungle Book* [CD]. Walt Disney Records.

Shore, C.M. (1995) *Individual differences in language development.* London [etc.]: Sage.

Silverman, B.W. (1981) 'Using Kernel Density Estimates to investigate Multimodality.' *Journal of the Royal Statistical Society. Series B, Methodological* **42**, 97-99.

Silverman, B.W. (1983) 'Some properties of a test for multimodality based on kernel density estimates.' In J.F.C. Kingman & G.E.H. Reuter (eds.) *Probability, statitics and Analysis.* Cambridge: CUP.

Sinha, C., Thorseng, L.A., Hayashi, M. & Plunkett, K. (1994) 'Comparative spatial semantics and language acquisition: evidence from Danish, English, and Japanese.' *Journal of semantics* **11**, 253-287.

Sternberg, R.J. & Okagaki, L. (1989) 'Continuity and discontinuity in intellectual development are not a matter of 'either-or'.' *Human development* **32**, 158-166.

Stewart, I. (1989) *Does God play dice? The mathematics of Chaos.* Cambridge (Mass.): Blackwell.

Stewart, I.N. & Peregoy, P.L. (1983) 'Catastrophe theory modeling in psychology.' *Psychological bulletin* **94** (2), 336-362.

Störig, H. J. (1972) *Geschiedenis van de filosofie [History of philosophy].* Utrecht: Het Spectrum. (Translation of *Kleine Weltgeschichte der Philosophie*, Stuttgart: Kohlhammer-Verlag.)

Sugarman, S. (1983) 'Discussion: empirical versus logical issues in the transition from prelinguistic to linguistic communication.' In R.M. Golinkoff (ed.) *The transition from prelinguistic to linguistic communication.* Hillsdale, NJ.: Lawrence Erlbaum Associates.

Sussmann, H.J. & Zahler, R.S. (1978) 'A critique of applied catastrophe theory in the behavioral sciences.' *Behavioral science* **23**, 383-389.

Ta'eed, L.K., Ta'eed, O. & Wright, J.E. (1988) 'Determinants involved in the perception of the Necker cube: an application of catastrophe theory.' *Behavioral science* **33**, 97-115.

Tesser, A. & Achee, J. (1994) 'Agression, love, conformity and other social psychological catastrophes'. In R.R. Vallacher & A. Nowak (eds.) *Dynamical systems in social psychology*. San Diego, Academic Press.

Thatcher, R.W. (1991) 'Maturation of the human frontal lobes: physiological evidence for staging.' *Developmental neuropsychologica* **7** (3), 397-419.

Thelen, E. (1989) 'Self-organization in developmental processes: Can systems approaches work?' In M.R. Gunnar & E. Thelen (eds.) *Systems and development.* Hillsdale, N.J.: Erlbaum.

Thelen, E. & Ulrich, B.D. (1991) 'Hidden skills: A dynamic systems analysis of treadmill stepping during the first year.' *Monographs of the society for research in child development* **223**, 56, 1.

Thelen, E. & Smith, L.B. (1994) *A dynamic systems approach to the development of cognition and action.* Cambridge (Mass.): MIT-press.

Thom, R (1975) *Structural stability and morphogenesis:an outline of a general theory of models*. Reading, Mass.: Benjamin.

Vainikka, A. (1994) 'Case in the development of English Syntax.' *Language Acquisition* **3**, 257-325.

Valian, V. (1991) 'Syntactic subjects in the early speech of American and Italian children.'*Cognition* **40**, 21-81.

Van der Maas, H.L.J. (1993) *Catastrophe analysis of stagewise cognitive development: model, method and applications.* Dissertation, University of Amsterdam.

Van der Maas, H.L.J. & Molenaar, P.C.M. (1992) 'Stagewise Cognitive Development: An Application of Catastrophe Theory.' *Psychological Review* **99**, 3, 393-417.

Van der Stelt, J.M. (1994) *Finally a word.* Amsterdam: IFOTT.

Van Geert, P. (1986a) *Theory building in developmental psychology.* Amsterdam: North Holland.

Van Geert, P. (1986b) 'The concept of development.' In P. van Geert (ed.) *Theory building in developmental psychology.* Amsterdam: North Holland.

Van Geert, P. (1986c) 'The structure of developmental theories.' In P. van Geert (ed.) *Theory building in developmental psychology.* Amsterdam: North Holland.

Van Geert, P. (1991) ‘A dynamic systems model of cognitive and language growth.’ *Psychological review* **98** (1), 3-33.

Van Geert, P. (1994a) *Dynamic systems of development: change between complexity and chaos.* New York: Harvester Wheatsheaf.

Van Geert, P. (1994b) ‘Vygotskian dynamics of development.’ *Human development* **37**, 346-365.

Van Geert, P. (1995) ‘Dimensions of change: a semantic and mathematical analysis of learning and development.’ *Human Development* **38**, 322 - 331.

Verrips, M (1991) ‘Models of development.’ In M. Rothweiler (ed.) *Linguistische Berichte (Sonderheft Grammatikerwerb)*, 11-21.

Villiers, J. de (1992) ‘On the acquisition of functional categories: a general commentary.’ In J.M. Meisel (ed.) *The acquisition of verb placement, functional categories and V2 phenomena in language development.* Dordrecht, Kluwer.

Wal, S. van der (1996) *Negative Polarity Items & Negation: Tandem Acquisition*. PhD. Dissertation, University of Groningen. (Grodil series, no. 17)

Wanner, E. & Gleitman, L.R. (eds.) (1982) *Language acquisition: the state of the art.* Cambridge, Cambridge University Press.

Weerth, C. de & van Geert, P. (1998) ‘Periods of instability associated to infantile developmental transitions.’ In press. *British journal of developmental psychology.*

Weissenborn, J., Goodluck, H. & Roeper, T. (1992) ‘Introduction: old and new problems in the study of language acquisition.’ In J. Weissenborn, H. Goodluck & T. Roeper (eds.) *Theoretical issues in language acquisition: continuity and change in development.* Hillsdale, N.J. [etc]: Erlbaum.

Weissenborn, J., Goodluck, H. & Roeper, T. (eds.) (1992) *Theoretical issues in language acquisition: continuity and change in development.* Hillsdale, N.J. [etc]: Erlbaum.

Wells, G. (1985) *Language development in the pre-school years.* Cambridge [et al.]: Cambridge University Press.

White, L. (1982) *Grammatical theory and language acquisition.* Dordrecht, Foris.

Wijk, C. Van & Kempen, G. (1980) ‘Funktiewoorden - een inventarisatie voor het Nederlands: an inventory of Dutch function words.’ *ITL: review of applied linguistics* (47), 52-68.

Wijnen, F. (1990) *On the development of language production mechanisms.* Dissertation University Nijmegen.

Wijnen, F., Krikhaar, E. & Os, E. den (1994) ‘The (non)realization of unstressed elements in children's utterances: evidence for a rhythmic constraint.’ *Journal of child language* **21**, 59-83.

Wijnen, F. (1995) ‘Clause structure develops.’ In M. Verrips & F. Wijnen (eds.) *Papers from the Dutch-German Colloquium on Language Acquisition* (Amsterdam Series in Child Language Development 3).

Wijnen, F. (1996) ‘Incremental acquisition of phrase structure.’ In J.N. Beckman (ed.) *Proceedings of NELS 25, Vol 2. Amherst, MA: GLSA Publications*, 105-118.

Wijnen, F. (1997a). ‘Temporal reference and eventivity in root infinitivals.’ In J. Schaeffer (ed.) The Interpretation of Root Infinitives and Bare Nouns in Child Language. *MIT Occasional Papers in Linguistics* **12**, 1-25. Cambridge, MA: MITWPL.

Wijnen, F. (1997b). ‘Functionele categorieën in Nederlandse kindertaal.’ *Nederlandse Taalkunde* **3**. In press.

Wimmers, R.H., Savelsbergh, G.J.P., Beek, P.J. & Hopkins, B. (1998) ‘Transitions in the development of eye-hand coordination.’ In press. *British Journal of Developmental Psychology*.

Wimmers, R. (1996) *Grasping developmental change: Theory, Methodology and Data.* Enschede: Printpartners Ipskamp.

Zeeman, E.C. (1976) ‘Catastrophe theory.’ *Scientific American* **234** (4), 65-83.

Zwart, J.W. (1993) *Dutch Syntax: a minimalist approach.* PhD. Dissertation, University of Groningen. (Grodil series, no. 10).

Appendix

Catastrophic equations and control variables

Control variables create a bifurcation set (in the Cusp in figure 3.1, these control variables are constructed by the y-axis and z-axis), and these variables control the dependent variables in the following fashion.

$$(I)\ \alpha = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + ... + \alpha_n x_n$$
$$(II)\ \beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + \beta_n x_n$$

Control variables α and β control a developmental process. These equations are used in Cobb's program. Furthermore, these equations of the control parameters are build up from several elements (α_1 to α_n and β_1 to β_n).

So far, it is unclear what the control parameters are in language development. Finding control parameters might strengthen the claim that the introduction of function words is catastrophic. The use of equations I and II is useless till control parameters have been found. Therefore, these equations could not be used in this study.

List of function words in the analyses:

Articles	**Modals** (finite forms, past tense, participium)	**Pronouns**
de een het &6 &t &6n &6t 't 'n d'	**kunnen** (kan, kunt, kun, kon, konden, gekund) **willen** (wil, wilt, wilde, wou, wilden gewild) **moeten** (moet, moest, moesten, gemoeten) **mogen** (mag, mocht, mochten, gemogen) **hoeven** (hoef, hoeft, hoefde hoefden, gehoeven) **zullen** (zal, zult, zou) **blijken** (blijk, blijkt, bleek, bleken, gebleken) **lijken** (lijk, lijkt, leek, leken, geleken) **schijnen** (schijn, schijnt, scheen, schenen, geschenen)	1. Determinative dit, ditte, deze, die, dat, datte 2. Possesive mijn, m'n, je, jouw, zijn, z'n, haar, d'r, 'r, ons, onze, uw, jullie, hun 3. Personal a. Nominative Single / Plural ik, 'k, jij, je, hij, zij, ze, het, 't, men / wij, we, jullie, u, gij, ge, zij b. Dative & Accusative Single / Plural me, jou, ie, hem, 'm, haar, 'r / ons, hen, hun

Summary in Dutch / Nederlandse samenvatting

Dit proefschrift behandelt een probleem, dat oud als de Griekse cultuur is. De vraag, bestaat er verandering in de natuur, houdt de gemoederen tot op heden nog steeds bezig. Destijds, in de Griekse oudheid, was het antwoord op de vraag tweeledig. Terwijl Parmenides stelde dat de natuur statisch was, veronderstelde Heraclites juist dat alles in de natuur aan verandering onderhevig is. Verandering is belangrijk in de ontwikkelingspsychologie, dat wil zeggen op het niveau van de menselijke ontwikkeling (de ontogenesis). Verandering is eveneens het onderwerp van dit proefschrift. In het eerste hoofdstuk staat de centrale vraagstelling vermeld, en deze luidt: wat zijn de paden van taalontwikkeling, en hoe zijn deze paden te verklaren aan de hand van psychologische en/of linguïstische theorieën? Deze vraag behelst het probleem van de studie van kwantitatieve taal*ontwikkeling*. Ontwikkeling kent diverse betekenissen, maar op het vlak van de taalstructuur is het begrip, in termen van kwantitatieve verandering, van ondergeschikt belang (in termen van aantallen studies) geweest aan linguïstische benaderingen.

In hoofdstuk 2 zijn twee begrippen cruciaal, namelijk taal en ontwikkeling. Om met de laatste te beginnen: theoretische overwegingen hebben ertoe geleid dat drie vragen in de ontwikkelingspsychologie steeds terug komen. Daar komt bij dat deze vragen niet afdoende beantwoord worden. Deze vragen over *oorsprong*, *nieuwheid*, en de *vorm van verandering* worden in de eerste paragraaf behandeld aan de hand van ontwikkelingstheorieën die samen zijn te vatten in *stadiumgewijze* en *stadiumloze* theorieën. Alle theorieën moeten bespreken hoe iets van niets komt. Anders gezegd: hoe wordt verandering verklaard? Eveneens van belang is de mate waarin een uitspraak over (de ontwikkeling van) taalstructuur wordt gedaan. Taalontwikkeling in ontwikkelingstheorieën is in termen van de structuur van taal matig uitgewerkt (d.i. voornamelijk in termen van leren): er ontbreekt een model en een theorie van de grammatica van een taal.

Na een korte bespreking van het *nativisme-empirisme* debat ga ik in op de discussie *discontinuïteit en continuïteit*, met als resultaat een chaos aan definities. In de psychologie zijn deze begrippen niet eenduidig omschreven. Dit kan evenmin van het begrip stadium worden gezegd. Stadia worden in verband gebracht met (dis)continuïteiten. In het laatste deel van deze paragraaf ga ik in op transities, d.i. snelle overgangen waarbij nieuwe structuren in de ontwikkeling van gedrag ontstaan. De mechanismen die hierbij een rol spelen worden besproken, en voorbeelden van onder andere motorische en cognitieve ontwikkelingsstudies waarin men dit soort overgangen heeft bestudeerd.

In paragraaf 2 van dit hoofdstuk wordt het systeem *taal* besproken, met name principes die in de linguïstiek inmiddels gemeengoed geworden zijn. Een verklaring van wat taal is wordt opgehangen aan (bijvoorbeeld) dieptestructuren uit de theorie van de generatieve grammatica. De waarschijnlijk grootste winst van de linguïstiek is de verklaring waarom bepaalde structuren in een taal wel en andere niet mogelijk zijn, en de verklaring voor het verband tussen talen en de ordening van woorden in een taal. Dit verband wordt weergegeven met zogenaamde syntactische bomen die de uitdrukking van een formele en universele taaltheorie zijn.

De paragraaf over taalontwikkeling gaat over de voorspellingen van taalontwikkeling. Daar komt de Universele Grammatica (UG) aan de orde, een van de grootste winsten van de linguïstiek. Deze aan alle talen onderliggende grammatica bestaat uit een aantal basale principes dat voor alle talen geldt. De UG vormt eveneens de link met het begrip discontinuïteit. Dat wil zeggen, discontinuïteit in de linguïstiek houdt in dat er een schending van UG is. Evenals de definities van stadia worden verschillen en overeenkomsten met definities van de (ontwikkelings)psychologie genoemd. De conclusie luidt dat de UG geen kwantitatieve voorspellingen voor ontwikkeling geeft.

In de laatste paragraaf, § 3, zet ik alles nog eens op een rijtje, waarbij een van de conclusies luidt dat de begrippen discontinuïteit en stadia vaag en multi-interpretabel zijn, en bovendien verschillen tussen de diverse vakgebieden. De vraag is ook hoe de definities uit de literatuur op empirische data moeten worden toegepast. Een mogelijke oplossing wordt aangereikt in hoofdstuk 3.

Hoofdstuk 3 bevat een inleiding in de wiskunde van de niet-lineaire dynamische systemen, met als doel twee theorieën en modellen te presenteren. Enerzijds is dat de Dynamische Systeem Theorie (DST), anderzijds is dat de Catastrofe Theorie (CT). Om te beginnen met de eerste soort modellen: DST is een tak van de wiskunde die poogt op basis van niet-lineaire vergelijkingen en op basis van continuïteit een verklaring voor processen te geven. Vanaf het eerste niveau, de initiële groeitoestand, groeit een systeem naar een eindtoestand. Die toestand wordt de Carrying Capacity genoemd (letterlijk: draagcapaciteit). Naast de begin- en eindtoestand is er nog een parameter, namelijk de groeiparameter, die de snelheid van ontwikkelen bepaalt. De huidige toestand wordt gegenereerd door middel van (o.a.) de output van de vorige toestand. Dit wordt iteratie genoemd, die uitdrukt dat tijdsmomenten niet onafhankelijk van elkaar zijn.

CT is een theorie die verklaart hoe plotselinge evenwichtsveranderingen plaatsvinden. In tegenstelling tot DST beschrijft CT discontinuïteit. Acht indicatoren, Catastrofevlaggen genoemd, kunnen gezamenlijk onderscheid maken tussen een werkelijke transitie (een kwalitatieve verandering in termen van reorganisatie en instabiliteit) en een verandering die continu is. Door naast een analyse van spontane data experimenteel controle variabelen te manipuleren kan extra evidentie worden gevonden voor discontinuïteit.

De toepassing van CT (bijvoorbeeld in de ontwikkelingspsychologie) is niet onomstreden, mede vanwege het karakter van menselijk gedrag en de problematiek van de statistiek van CT. Ondanks deze problemen zijn er in de afgelopen jaren aanwijzingen gevonden dat sommige processen, zoals motorische ontwikkeling en conservatie, niet lineair en zelfs discontinu zijn.

Hoofdstuk 4, over de methode, staat in het teken van het werk van met name Brown die veel aandacht heeft besteed aan de ontwikkeling van functiewoorden (FW). FW vormen een gesloten klasse van woorden. Het is een klasse waarin geen woorden (d.w.z. types) bijkomen in de sociogenetische ontwikkeling van een taal. Andere kenmerken zijn dat ze hoog frequent in een taal zijn, weinig klemtoon in een zin bezitten, en een systematisch vaste plaats in een zin (zo komen lidwoorden voor zelfstandige naamwoorden) kennen, en al met al syntactisch zeer belangrijk zijn. De ontwikkeling van functiewoorden in de taal van een kind heeft geleid tot de conclusie “er is sprake van een vlugge ontwikkeling van telegramstijl naar differentiatiestijl", hetgeen een magere conclusie is te noemen als het gaat om onderscheid tussen continuïteit en discontinuïteit te maken.

In paragraaf 2 behandel ik de voorwaarden waaraan een onderzoek als het onderhavige moet voldoen. Voor de kwantificering van de data gebruik ik *frequentieanalyses* als een index voor het gebruik van een bepaalde grammaticale regel. Anders gesteld: hoe vaak komt een bepaald woord in de taal van een kind voor? Het gebruik van frequentieanalyses moet aan een aantal randvoorwaarden voldoen. Ten eerste moeten de situaties tijdens de opnames zoveel mogelijk gelijk zijn (representativiteit). Ten tweede moet de lengte van de opnames een minimum hebben, en de opnames moeten gelijk zijn in lengte. Ten derde moet het interval tussen de diverse opnames voldoende klein zijn. Het gebruik van ofwel absolute, ofwel proportionele frequentie scores is daarbij niet van wezenlijk belang.

De zes kinderen, een bespreking van het aantal opnames en transcripten (180 in totaal), en de programma's die zijn gebruikt voor het maken van frequentie analyses staan in de laatste paragraaf van dit hoofdstuk vermeld. Uiteindelijk zal de volgende vraag worden beantwoord: is er sprake van een discontinue ontwikkeling, gedefinieerd in Catastrofe-theoretische termen, of continue ontwikkeling van functiewoorden, in termen van Dynamische systeem theorie?

Het antwoord op die vraag volgt in hoofdstuk 5. In de eerste paragraaf presenteer ik aanwijzingen voor drie CT-vlaggen: plotselinge sprong, bimodaliteit, en ongewone variantie. Twee modaliteitsanalyses suggereren dat de data multimodaal verdeeld zijn. Bovendien wijst het Cuspfit programma in de richting van een beste fit van het cusp model. Echter, dit programma heeft als bezwaar dat curves van snelle groei ook een beste fit van de cusp zullen krijgen. Al met al is het bewijs van discontinuiteit (d.w.z. een catastrofe fasetransitie) niet overtuigend.

Daarom behandel ik in de tweede paragraaf continue modellen. Het groeimodel geeft een betere fit dan het lineaire. Simulaties van ontwikkeling op basis van een lineaire vergelijking met ruis kunnen de gevonden empirische data niet reproduceren. Middeling van de data op leeftijd laat een graduele groei zien, terwijl bij middeling op de eerste substantiële toename een plotselinge verandering is waar te nemen.

Een verklaring voor de gevonden resultaten kan niet worden gevonden in de ontwikkeling van onafhankelijke variabelen als een groei van het aantal woorden per sessie, of een plotselinge toename van het aantal functiewoorden van de ouders.

In hoofdstuk 6 wordt geredeneerd dat de data niet eenduidig in de richting wijzen van discontinuïteit, en dat dus niet gekozen kan worden voor een catastrofe-model (bijvoorbeeld de cusp): te weinig kinderen vertonen te weinig vlaggen. Naar aanleiding van de diverse verklaringen ga ik dieper in op wat de data betekenen voor de studie van taalontwikkeling. Er zijn diverse verklaringen mogelijk. Ten eerste kan de snelle verandering zijn veroorzaakt door de syntactische ontwikkeling van een kind: doordat de taal van een kind in termen van finiete werkwoorden en langere zinnen is veranderd zijn er *slots* ontstaan voor functionele categorieën. Ten tweede is er de mogelijkheid dat er een plotselinge neurologische ontwikkeling een reden voor snelle taalontwikkeling is, maar de relatie tussen neurologisch ontwikkeling en taalontwikkeling is nauwelijks onderzocht. Neurologische verandering biedt een correlaat van taalontwikkeling, geen verklaringsgrond. Andere verklaringen (bijvoorbeeld een conflict van regels) bieden geen oplossing van het optreden van de plotselinge sprong.

Kwantitatieve verklaringen voor de gevonden data zijn een dynamische systeem verklaring die ondersteund wordt door de ontwikkeling van het geheugen aan de ene kant, en een linguïstische verklaring aan de andere kant. Dat geheugen wordt steeds beter in het vasthouden van (o.a.) onbeklemtoonde woorden als functiewoorden. De taalkundige ontwikkeling geeft ruimte voor functiewoorden doordat zinnen langer worden, en werkwoorden hun plaats in de zin hebben veroverd, waarna syntactische openingen voor functiewoorden zijn ontstaan. Bovendien worden de gevonden data onderbouwd door instabiliteit in het zich ontwikkelende taalsysteem. Het begrip loopt voor op de productie bij de ontwikkeling van functiewoorden, en fonologische beperkingen verdwijnen geleidelijk.

Een dynamisch model van taalontwikkeling biedt een verklaring voor taalontwikkeling. Met behulp van psychologische en linguïstische inzichten beschrijft dit model de verwerving van de woordenschat tot de constructie van een zin. Maar behalve beschrijvend is dit model ook predictief. Als het model correct is, dan betekent een verandering van parameter waarden (bijvoorbeeld de groeiratio) dat taalontwikkeling een ander verloop zal en moet nemen. Het geeft echter geen antwoord op de vraag van nieuwheid, een vraag die direct gerelateerd is aan verandering. Daarom is ook voorgesteld dat dynamische systemen niet op zichzelf staand door het leven gaan, maar dat ze een verstandshuwelijk met connectionistische modellen aangaan. Deze laatste modellen verklaren hoe, zonder de aanname van innateness (een van de grootste discussiepunten en struikelblokken in de studie van taalontwikkeling), patronen in taalontwikkeling ontstaan, terwijl de dynamische modellen de patronen van ontwikkeling modelleren, en veranderingen verklaren.

De data wijzen in de richting van op zijn minst een zeer snelle taalontwikkeling. Deze bevindingen zullen gevolgen hebben voor de linguïstische studie van de ontwikkeling van taal. Immers, een theorie van taalontwikkeling moet naast een structurele verantwoording het proces beschrijven en verklaren. Ik heb beargumenteerd dat beide beschrijvingen niet onafhankelijk van elkaar kunnen bestaan als de psychologische ontwikkeling van linguïstische taalstructu(u)r(en) dient te worden begrepen. Daarbij is inzicht in variabiliteit belangrijk voor een compleet inzicht in ontwikkelingsprocessen.

List of publications

Papers

Cats, M. & Ruhland, R. (In press) ‘Simulatie van taalontwikkeling. [Simulation of language development]’ To appear in *Tijdschrift voor Informatica en Maatschappij* (TIM).

Ruhland, R., Wijnen, F. & van Geert, P. (1995) ‘An exploration into the application of dynamic systems modelling to language acquisition.’ In F. Wijnen and M. Verrips (eds.) *Amsterdam Series in Child Language Development: approaches to parameter setting* 4, 107 - 134.

Ruhland, R. (1996) ‘Frequency analysis in language development.’ In R. Jonkers, E. Kaan & A. Wiegel (eds.) *Language and Cognition 5 (Yearbook 1995 of the research group for Theoretical and Experimental Linguistics of the University of Groningen)*. University of Groningen, Dept. of Linguistics.

Ruhland, R. (1997) ‘Closed class words in motion.’ To appear in the proceedings of the KNAW-seminar *Non-linear approaches to development*.

Ruhland, R. & van Geert, P. (1997) ‘Jumping into syntax: transitions in the development of closed class words.’ In press. *British journal of developmental psychology*.

Posters

‘Sudden changes in development: Catastrophe theory in Languageland.’ Poster presented at the *Groningen Assembly on Language Acquisition*. Held at the University of Groningen, The Netherlands. September 7-9 1995.

Abstracts

‘Transities in de vroege taalontwikkeling.’ Psychon *Ontwikkelingspsychologie* (Congres of the Dutch Developmental Psychology Society), March 18-19 1993. Dalfsen, The Netherlands.

‘The development of function words.’*Dutch-German Colloquium on Language Acquisition*, September 1-2 1994. Groningen, The Netherlands.

‘Transities in de vroege taalontwikkeling.’ SGW *Ontwikkelingspsychologie* (Congres of the Dutch Developmental Psychology Society), May 18-19 1995. Dalfsen, The Netherlands.

‘Closed class words in motion.’ KNAW symposium *Non-linear approaches to development*, Januari 29-31 1997. Amsterdam, The Netherlands.

‘Transities in functiewoorden.’ SGW *Ontwikkelingspsychologie* (Congres of the Dutch Developmental Psychology Society), April 17-18 1997. Leusden, The Netherlands.

Review

The Acquisition Device smiles. Review of *Syntactic development* by William D. O’Grady (1997), Chicago [etc.]: The university of Chicago Press. 409 p. ISBN 0-226-62077-8 (pbk.). To appear in *Early Development and Parenting*.

ᚷᛟᛁᚾᚷ ᛏᚻᛖ ᛞᛁᛋᛏᚪᚾᚳᛖ

ᚪ ᚾᛟᚾ-ᛚᛁᚾᛖᚪᚱ ᚪᛈᛈᚱᛟᚪᚳᚻ ᛏᛟ ᚳᚻᚪᚾᚷᛖ
ᛁᚾ ᛚᚪᚾᚷᚢᚪᚷᛖ ᛞᛖᚹᛖᛚᛟᛈᛗᛖᚾᛏ

ᚻᚷ ᛋᚢᚢᛚᚪᚾᛞ

Stellingen

behorend bij het proefschrift

Going the distance: a non-linear approach to change in language development.

1. In plaats van leeftijd kan in de studie naar taalontwikkeling beter gesproken worden van taaltijd. *(Dit proefschrift)*

2. Het gebruik van groepsgemiddelden als basis voor een model van taalontwikkeling berust op misverstanden. *(Dit proefschrift)*

3. Linguïstische modellen lenen zich niet voor een kwantitatieve, dynamische benadering van taal. *(Dit proefschrift)*

4. De ontwikkeling van kindertaal is, in haar meest letterlijke betekenis en gemeten aan de hand van de syntaxis, een kwestie van continuïteit. *(Dit proefschrift)*

5. De toepassing van het begrippenpaar continuïteit-discontinuïteit in diverse disciplines leidt tot verwarring en onbegrip. *(Dit proefschrift)*

6. Kunst is een kwestie van kijken, luisteren en een eigen visie toevoegen, wetenschap is een kwestie van zien, vragen en de eigen mening weglaten.

7. Liefde infantiliseert.

8. Het leven is het mooist als het voorbij de onwennigheid is, maar voor de gewenning blijft.

9. Als geweld tegenover liefde staat, staat zinloos geweld tegenover zinloze liefde.

10. Koketteren met gecijferdheid heeft als onmiddellijk gevolg dat elke vertelling als ‘getalmoeheid’ wordt aangerekend.

11. Wie aan de tweede kamerleden en de kiezers vraagt om achter de polarisatie in een politieke discussie te kijken, vraagt eigenlijk om te zien wat er niet is. Het opvoeren van een discussie in termen van extreme standpunten is dan ook niet meer dan effectbejag en slecht theater.

12. De perfecte balans tussen sport en kunst is te vinden in de NBA-basketbalcompetitie.

13. Een Groningse stelling van een proefschrift is gedocumenteerde wijsheid en nieuws als de Loeks er gewag van maakt.

14. 5 jaar in de wetenschap staan in geen verhouding tot 3 kwartier in de televisiequiz *2 voor 12*, noch in financieel, noch in intellectueel opzicht. De suggestie, gedaan in Veenstra (1995), dat elke aio en oio goed geld kan verdienen met televisiespelletjes, moet in het geval van *2 voor 1*2 van de hand gewezen worden, aangezien velen niet veel weten, niet weten waar ze het moeten zoeken, laat staan weten waar ze het moeten opzoeken.

    J. Veenstra (1995) *Attention in preschool children with and without signs of ADHD*. Dissertation, University of Groningen.

15. In a period of transition timid elements will run for cover.

    The Hudsucker Proxy (1994) Directors: Joel & Ethan Coen.

16. De Overijsselse IJssel stroomt op majestueuze wijze langs de Veluwe. Uiteindelijk, voorbij Zwolle, leidt dat tot een *menage a trois* met de Vecht en het Zwarte Water. Die drie-eenheid toont aan dat het paradijs niet ver weg is.

Rick Ruhland
Groningen, 1998